M 2014

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Integrating RNA Polymerase II dynamics with alternative polyadenylation

JOANA TROÇA CORREIA MAIA

DISSERTAÇÃO DE MESTRADO APRESENTADA
À FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO EM
BIOENGENHARIA

- This page was intentionally left blank. -

U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

Integrating RNA Polymerase II dynamics with alternative polyadenylation

Joana Troça Correia Maia

Master Thesis

Master in Bioengineering - Molecular Biotechnology
2013/2014

Supervisor: Alexandra Moreira

Porto, September 2014

Abstract

In eukaryotes, gene expression is regulated at different levels, including transcription and RNA processing. Alternative polyadenylation in the 3'untranslated region during pre-mRNA processing has a fundamental role in different biological processes and disease.

In this thesis, using the genes *polo* and *CG6024* as a model, the impact of an increased RNA polymerase II elongation rate on alternative polyadenylation is characterized in a *D. melanogaster* mutant (RpII140) that has 50% faster transcription elongation rate. The model genes contain two pA signals, with opposite efficiency strengths, which are used to produce two mRNA isoforms with different 3'UTR lengths. It was observed a decrease of total and pA1 mRNA levels and different results regarding pA2 mRNA levels in both genes, in this mutant fly. These results suggest that faster RNA polymerase II elongation rate reduces the 3'end formation efficiency.

In addition, *polo* poly(A) site usage was determined when the elongation factors dELL and dSpt6 were depleted via RNAi. It was observed that dELL enhances the usage of *polo* proximal poly(A) site. The results for dSpt6 were inconclusive due to variations in the expression of the reference gene used (*7SL*).

The work described in this thesis shows an important role of the RNA polymerase II elongation rate in alternative polyadenylation and also suggests a role of dELL in poly(A) site choice.

Keywords: Alternative Polyadenylation; Elongation; Transcription; *Drosophila melanogaster*; Pol II elongation rate; dELL; dSpt6; mRNA isoforms; 3'UTR

- This page was intentionally left blank. -

Acknowledgments

It is a truth universally acknowledged, that to attain our goals individual commitment is not enough. Only with the guidance, assistance and collaboration from others, extraordinary achievements can be accomplished. For that reason, I am especially grateful:

To *Alexandra Moreira*, for having accepted me in her lab and for do much more than teach me. Thank you for encouraged, instructed, mentored, guided and more important inspired me.

To *Jaime Freitas*, for the help, for transmitting me so many knowledge and enlightening my doubts. Thank you for teach me so many things and share with me what you know, with that you opened my eyes, mind and my heart to the scientific world.

To *Mafalda Pinho*, *Inês Baldaia*, *Ana Curinha*, *Sandra Braz*, *Inês Boal*, *Marta Oliveira*, *Isabel Castro* and *Andrea Cruz*, for the companionship in the lab and the assistance provide in this journey.

To *Maria Pinto*, *Sara Rocha*, *Joana Costa* and *Francisco Coelho*, for always understanding me and being so good listeners, for all the good advices and laughs…great thanks! I could not have asked for better companions for this journey.

To *my "Biotech" friends from the university*, for the fellowship and all the good moments we spent together.

To *Nuno Matos*, *Sílvia Godinho*, *Sara Marques, Maria Pinho* and *Luís Santos*, for the friendship and continuous support over all these years.

To *Fábio*, for all the support, the patience, the intelligent advices and continuous companionship and love. I don't have enough words to describe all my gratitude

À minha família, pelo apoio incondicional, por me encorajarem nos momentos mais difíceis e por saberem lidar tão bem com a minha má disposição nessas alturas. Obrigada por me concederem a oportunidade de tirar este curso, obrigada por tudo. A vós devo tudo o que hoje sou…Obrigada!

- This page was intentionally left blank. -

Table of Contents

List of Figures

- This page was intentionally left blank. –

List of Tables

- This page was intentionally left blank. –

List of Appendices

- This page was intentionally left blank. -

Abbreviations

Abd-B	Abdominal-B
ChIP	Chromatin Immunoprecipitation
Co-IP	chromatin -immunoprecipitation
CTD	C-terminal domain
DMSO	Dimethyl sulfoxide
DSIF	DRB sensitivity-inducing factor
dsRNA	double-stranded RNA
EDTA	[[2-(Bis-carboxymethyl-amino)-ethyl]-carboxymethyl-amino] acetic acid
ELL	eleven-nineteen lysine-rich leukemia
FBS	Fetal bovine serum
GTFs	general transcription factors
LEC	little elongation complex
mRNA	messenger RNA
NELF	negative elongation factor
Nt	nucleotides
Paf	Polymerase Associated Factor complex
PIC	pre-initiation complex
Pol II	RNA polymerase II
P-TEFb	positive transcription elongation factor
RNAi	RNA interference
s.e.m.	standard error of the mean
S2 cells	Drosophila Schneider cells
SEC	super elongation complex
Spt	Suppressor of Ty
TAE	Tris-acetate-EDTA
Ubx	Ultrabithorax
UTR	untranslated region

- This page was intentionally left blank. -

Chapter 1

Introduction

In eukaryotic cells, the molecular processes involved in messenger RNA (mRNA) production are certainly among the most complex events in cellular biology and biochemistry. In fact, it is not only necessary a diverse set of factors to start RNA synthesis and the elongation of the RNA chain, but also the accurate processing of the primary RNA transcript is extremely complex (Darnell 2013).

In the past, pre-mRNA processing events were thought to occur independently, until several experiments revealed that these mechanisms are tightly coupled to RNA polymerase II (Pol II) transcription, i.e. they are mostly co-transcriptional (Bentley 2014). Currently, the term co-transcriptional has come to imply a functionally significant coupling between transcription and RNA processing events, and since these processes are involved in the fine-tuning of gene expression, a great interest has been arising to understand the interplay between these two events.

1. Elongation: a crucial point not only in the transcription cycle

It is well established that transcription is one of the most highly regulated phenomenon in eukaryotic gene expression. This process is carried out by orthologous enzymes, multi-subunit DNA-dependent RNA polymerases (RNAPs) and essentially is comprised by three phases that are mechanistically distinct: initiation, elongation and termination (Svetlov & Nudler 2013; Palangat & Larson 2012). The first stage consists in the promoter recognition and opening of the DNA helix. In the next phase RNAPs extends the RNA transcript, and in the termination stage, dissociation of the RNA–DNA hybrid occurs with the nascent transcript being released from the DNA template (Nechaev & Adelman 2011; Palangat & Larson 2012). The cycle of eukaryotic transcription by Pol II is a complex and regulated event. Unquestionably, it is the coordinated

action of regulatory factors that assures the transcriptional competence of Pol II at different phases (Kwak & Lis 2013).

Transcription is a precisely timed event. It starts with the recruitment of RNA polymerase and general transcription factors (GTFs) to a target gene promoter. The recognition of the promoter begins with a sequence-specific binding of an activator to the enhancer DNA elements that stimulates the assembly of a large protein complex containing Pol II and multiple GTFs (TFIIB, TFIID, TFIIE, TFIIF and TFIIH) known as pre-initiation complex (PIC). This PIC undergoes a series of transformations as the nascent RNA is synthesized, and the key modification is the local melting of the DNA around the transcription start site (TSS) in which a double stranded promoter transforms to an single stranded promoter complex. The subsequent events as *de novo* RNA synthesis (formation of the first phosphodiester bond), abortive initiation (in which short transcripts are continually synthesized and released by polymerase still associated with the promoter), and promoter escape (a stage characterized by physical and functional instability of the transcription complex) precedes transcript elongation (Liu et al. 2013; Nechaev & Adelman 2011). During transcription initiation the initial length of the nascent transcript is crucial, since transcripts of less than 5 nucleotides (nt) are unstable (which results in a high frequency of abortive initiation), transcripts with 10 nt display a preference for promoter escape over abortive initiation and transcript with length of around 25 nt achieve productive initiation and consequently transcription elongation initiates (Saunders et al. 2006).

At this point, the initiation-elongation transition starts with the dissociation of initiation factors and recruitment of elongation factors to the Pol II transcription complex (Lidschreiber et al. 2013; Zhou et al. 2013). Elongation extends nascent RNA by one nucleoside monophosphate with pyrophosphate release, followed by Pol II translocation along the template by one nucleotide (Svetlov & Nudler 2013). Interestingly, in metazoans, Pol II elongation throughout a gene is not always a uniform event, since after transcription of the first ~20–60 nts Pol II pauses. This is a crucial step for approximately half of all active *Drosophila* and mammalian genes because the progress beyond this point is rate limiting. Therefore, many authors consider elongation as a process composed by two distinct stages: early elongation and productive elongation (Kwak & Lis 2013).

1.1. Pausing *vs* Productive elongation

Pol II is able to transit between promoter escape and fully productive elongation and often this event is complemented with an intermediate step of promoter-proximal pausing which is known as early elongation. After initial elongation of nascent RNA (up to ~10 nts), Pol II is capable of escaping the initiation stage. At this point GTFs such as TFIIB dissociate, however parts of the initiation complex may remain connected to the promoter in order to provide a scaffold for re-initiation. Once the transcript is longer than 12 nts and TFIIB is removed, the Pol II elongation complex becomes stably engaged and the early elongation stage begins. During the

course of early elongation in higher eukaryotes, Pol II pauses after transcribing ~20–60 nts. Pol II is most likely to be found in an arrested state (backtracked position), but the enzyme can be rescued from that state by binding the TFIIS, which stimulates cleavage of protruded nascent RNA and realignment of the new pre-mRNA 3' end with the Pol II active site (Saunders et al. 2006; Kwak & Lis 2013; Nechaev & Adelman 2011). Additionally, *in vitro* studies demonstrate that the early elongation complex has a strong tendency to pause, arrest and terminate transcription, which highlights the inefficiency of the event. This feature of the early elongation complex can have multiple reasons, including negative action of some elongation factors that inhibit synthesis through the promoter-proximal region or the fact that the early elongation complex need to experience some conformational modifications in order to became fully stable and processive (Nechaev & Adelman 2011).

Pioneering studies revealed that many factors and specific promoter features are essential for full regulation of early elongation, disclosing another critical layer of transcription control. Indeed, the early elongation mechanisms are connected to the promoter structure, which affects both transcriptional initiation and pausing (Gilchrist et al. 2010; Kwak & Lis 2013). Two main factors regulate Pol II pausing, the DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF), through binding to and inhibiting Pol II function (Marshall & Price 1992; Mandal et al. 2004). Regarding the functional role of transcriptional pausing, recent works suggest that the ability of Poll to pause modulates transcription kinetics, output, or the coordination of gene activation. Moreover, the potential of this event as a fine-tuned mechanism of regulating gene expression levels in response to a changing environment has been discussed (Adelman et al. 2009; Boettiger & Levine 2009; Nechaev & Adelman 2011).

To enter the stage of productive elongation Pol II needs to escape from early elongation which requires a dynamic transition from the paused state into its subsequent reinstatement in a productive mode (Kwak & Lis 2013). This shifting requires a highly controlled exchange of factors that is orchestrated mainly through phosphorylation of the C-terminal domain (CTD) of the largest Pol II subunit (Rpb1). This domain consists of tandem heptapeptide repeats (27–52) of the consensus sequence, YSPTSPS. Each repeat bears three serine residues (Ser2, 5 and 7) and a threonine residue (Thr4) that undergo reversible phosphorylation by CTD-modifying enzymes at specific stages of transcription, in the so-called the "CTD code" (Buratowski 2005; Hintermair et al. 2012; Heidemann et al. 2013; Egloff et al. 2012; Hsin & Manley 2012). During the transition from transcription initiation to elongation, Pol II changes from a hypophosphorylated to a hyperphosphorylated form. It is noteworthy that the various CTD modifications can affect Pol II conformation and ability to associate with different factors which modulate various events, including pre-mRNA processing, in particular 3' end formation (McCracken et al. 1997; Saunders et al. 2006).

In order to accomplish the transition into productive elongation recruitment of the positive transcription elongation factor (P-TEFb) is necessary. In fact, P-TEFb is able to reverse

the negative influence of DSIF and NELF by phosphorylating the largest component of pausing factors subunits (Suppressor of Ty (Spt) 5). This event also stimulates the function of TFIIF in productive transcription. Furthermore, P-TEFb phosphorylates other targets such as Ser-2 CTD, which help transcription through chromatin and pre-mRNA processing (Moore & Proudfoot 2009; Saunders et al. 2006).

Even after pause escape, Pol II still must surmount some roadblocks to achieve a dynamic environment for productive transcription. Consequently some specific post-translational modifications on the CTD and on the N-terminal histone tails of the nucleosome are necessary (Shandilya & Roberts 2012). Indeed, the organization of eukaryotic DNA into chromatin severely affects Pol II progress into productive elongation or even transcription termination (Gilchrist et al. 2010; Mischo & Proudfoot 2013). In many cases histone modifiers such as Spt6 (that remove nucleosome barriers by interacting directly with histones H3 and H4 and help to render the DNA template accessible for productive elongation (Bortvin & Winston 1996).

The overall transcriptional elongation rate relies on the balance between positive and negative regulatory factors, as well as on the chromatin conformation (*Table 1*). It must be highlighted that some elongation factors are associated to Pol II and enhance the elongation rate essentially by suppressing pausing factors (Palangat & Larson 2012).

Elongin and eleven-nineteen lysine-rich leukemia gene (ELL) are well described families of proteins that are able to increase the catalytic rate of Pol II (Lin et al. 2013; Kwak & Lis 2013). Recently, Lin *et al.* identified the super elongation complex (SEC), comprising the ELL family of Pol II elongation factors (ELL1, 2 and 3) and the positive transcription elongation factor P-TEFb, as many of the most frequent mixed lineage leukemia translocation partners. (Lin et al. 2010). In fact, SEC was found to regulate the transcriptional elongation checkpoint control (TECC) stage, which is an extremely relevant step in transcription, since it refers to the regulation of Pol II assembly and release from the paused stage. The TECC stage is crucial for the regulation of gene expression during development and its misregulation is associated with various diseases, including cancer (Smith, Lin & Shilatifard 2011).

Overall, these examples illustrate how much remains to be discovered regarding the already known elongation factors, and also suggest that the repertoire of proteins that affect elongation has not yet been completely revealed.

The entire transcription cycle is coordinated with other essential events that lead to proper processing of RNA. Indeed, as mentioned before, the CTD is not only subject to a plethora of post-translational modifications, but also acts as a scaffold for the recruitment of several factors involved in distinct stages of transcription as well as in co-transcriptional processes such as pre-mRNA 5' capping, splicing and 3' end formation (Lidschreiber et al. 2013; Moore & Proudfoot 2009). Therefore, the interactions between transcription factors and processing machineries form a tight regulated network of cross-stimulatory connections with a role on pre-mRNA processing (Mapendano et al. 2010).

Table 1- List of the most relevant factors involved in transcription elongation

Class	Factor	Properties	References
General transcription factors	TFIID	Assists in promoter structure for pausing	(Shopland et al. 1995)
	TFIIF	Increases elongation rate	(Bengal et al. 1991; Tan et al. 1995)
	TFIIS	Rescues backtracked Pol II	(Bengal et al. 1991; Kettenberger et al. 2004)
Pausing factors	NELF	Stabilizes Pol II pausing	(Narita et al. 2003; Yamaguchi et al. 1999)
	DSIF	Stabilizes Pol II pausing and facilitates elongation	(He et al. 2011; Yamada et al. 2006)
Positive elongation factor	P-TEFb	Stimulates pause release by phosphorylating NELF, DSIF, and Pol II CTD	(Kim & Sharp 2001)
Processivity factors	Elongin	Alleviates pausing, increases Pol II rate	(Aso et al. 1995)
	ELL	Alleviates pausing, increases Pol II rate	(Lin et al. 2013)
	SEC	Contains ELL and P-TEFb	(Lenasi & Barboric 2010)
Co-activator	Mediator	Recruits P-TEFb via SEC	(Takahashi et al. 2011)
Polymerase associated-complex	PAF	Platform to elongation factors assembly (e.g SEC)	(He et al. 2011)
Histone tail modifiers	SET1	Methylates histone H3-K4, localized to promoter and coding region	(Krogan et al. 2003)
	SET2	Methylates histone H3-K6, localized to coding regions	(Krogan et al. 2003)
Histone chaperone	Spt6	H3-H4 chaperone and Tracks with Pol II	(Saunders et al. 2003)

1.2. Elongation as a regulator of gene expression

In the past it was generally assumed that gene expression regulation took place only *via* promoters and enhancers, i.e. at the transcription initiation step. Nevertheless, now it is generally recognized that gene expression can be regulated at all steps of the transcription cycle, with an emphasis at the level of transcription elongation (Svejstrup 2013; Proudfoot 2004). Therefore, it has been a significant paradigm shift in the transcription field, since the elongation event is now appreciated as significant regulator of gene expression.(Svejstrup 2013).

The relevance of transcriptional elongation as a regulator of gene expression has just recently been recognized. In fact, recent work demonstrates that transcription elongation can adjust the outcome of gene expression, since it modulates cellular RNA levels, essential for a diverse range of biological pathways, differentiation, development and survival of all cells and organisms (Jennings 2013; Kwak & Lis 2013; Smith & Shilatifard 2013; Svejstrup 2013). Furthermore, deregulation of elongation can lead to the development of certain diseases (Li & Green 1996).

As mentioned before, transcription elongation factors have different functions, such as modulating the catalytic properties and processivity of Pol II and assisting the progression of the enzyme through repressive chromatin. Besides that, it is also described that these factors can have an impact in gene expression by modulating the elongation stage (Saunders et al. 2006). Indeed, this additional role of elongation factors is illustrated by Dürr *et al.*' s work, which showed that elongation factors Spt4/Spt5 can modulate the expression of *Arabidopsis* genes transcribed by Pol II, particularly influencing the transcription of genes involved in auxin signalling (Dürr et al. 2014) . Furthermore, Chopra *et al.* presented genetic and molecular evidence that the *Drosophila* Hox genes are subject to an additional layer of regulation at the level of elongation, since the Pol II elongation factors Elongin-A and Cdk9 proved to be essential for optimal Ultrabithorax (Ubx) and Abdominal-B (Abd-B) expression (Chopra et al. 2009).

Additionally, elongation can regulate gene expression by a functional coupling with other molecular events. As previously stated, many pre-mRNA processing events occur co-transcriptionally, therefore is not surprising that some processing events are affected by transcription, namely by elongation (Howe 2002). This is the case for splicing, which was shown that it is functionally coupled to transcription by correlating Pol II elongation rate with the ability of splicing factors to recognize splice sites of various strengths (Dujardin et al. 2014). As an example, Robert *et al*. showed that in α-tropomyosin gene, slower elongation favours utilization of weaker upstream splicing sites (Roberts et al. 1998). Consequently, the coupling between elongation and processing events, such as splicing, help cells to modulate the levels of expressed proteins and/or produce multiple protein isoforms from one single transcriptional unit.

2. Alternative Polyadenylation in the 3' end formation

To obtain a mature mRNA the primary RNA transcript (pre-mRNA) needs to be processed. Pre-mRNA processing includes various steps that are mostly co-transcriptional: capping, editing, splicing and polyadenylation (Lutz & Moreira 2010; Yang & Doublié 2011; Neugebauer 2002).

During 5' end cap modification, the 5′ triphosphate of the pre-mRNA is cleaved and a guanosine monophosphate is added which is subsequently methylated (Lewis & Izaurralde 1997). The addition of the 5′ cap provide resistance of the mRNA molecule to 5′ to 3′ exonucleases, which expands the (pre)-mRNA lifetime. Additionally, the 5' cap interacts with protein factors involved in other cellular processes such as pre-mRNA splicing, nucleocytoplasmic RNA export and localization and translation (Cougot et al. 2004; Neugebauer 2002).

RNA editing is a modification of individual RNA residues which are converted to alternative bases. These base modifications are performed by adenosine deaminases acting on RNA (ADAR) that substitute adenosine to inosine (A-to-I editing), or by apolipoprotein B mRNA editing enzyme (APOBEC1), which converts cytosine with uracil. This process increases the complexity of the mRNA isoforms at the single-base level and consequently may result in an amino acid sequence change in the protein, in differences in alternative splicing decisions or in nuclear retention of the transcript (Dillman et al. 2013; Jin et al. 2007; Garncarz et al. 2013).

Another significant process in mRNA maturation is the precise removal of noncoding sequences (introns) and the joining of the remaining sequences (exons) in a multistep reaction, a process known as RNA splicing. This modification is mediated by a catalytically active form of a megacomplex named the spliceosome (Montecucco & Biamonti 2013; Dujardin et al. 2013). It is noteworthy that some introns can be retained and particular exons can be entirely or partially skipped, therefore more than one mRNA molecule from a single transcriptional unit can be produced by alternative splicing. These splicing events have the biological advantage of increasing the proteome diversity, since they allow the same transcriptional unit to produce many diverse proteins that may have different localizations, stabilities and functions (Dujardin et al. 2013; Alberts et al. 2008; Sanchez et al. 2011). In addition to the production of a new proteins, alternative splicing may shift the reading frame which may result in transcripts with premature stop codons (Lewis et al. 2003; Resch et al. 2004). Some splicing events were shown to be regulated in response to different cell stimuli which ensure the production of the right proportion of the different isoforms in the different occasions (Castro et al. 2007; da Glória et al. 2014). In fact, in order to guarantee plasticity, high specificity and fidelity, the alternative splicing process requires a very precise regulation (Nogués et al. 2003; Dujardin et al. 2013).

The generation of mature mRNAs also requires correct 3′ end processing. mRNA 3' end formation requires two tightly coupled steps involving an endonucleolytic cleavage of the nascent transcript, followed by the polymerization of the adenosine tail (polyadenylation reaction) (Rehfeld et al. 2013).

In eukaryotes, polyadenylation of transcripts generated by Pol II is an essential step of gene expression and also a versatile mechanism of gene regulation (Colgan & Manley 1997; Zhao & Hyman 1999; Millevoi & Vagner 2010; Shepard et al. 2011). Effective pre-mRNA 3' end cleavage requires a core molecular machinery that includes four multisubunit protein complexes: cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF),

cleavage factors Im and IIm (CFIm and CFIIm) (Gruber et al. 2013; Yang & Doublié 2011). At the 3' end of the upstream cleavage product, a poly(A) tail is added by nuclear poly(A) polymerases (PAP), α (PAPOLA), β (PAPOLB), or γ (PAPOLG). The length of the poly(A) tail is determined by the nuclear poly(A)-binding protein 1 (PABPN1), although the length greatly differ between species (Gruber et al. 2013; Wahle 1991; Wahle & Keller 1996; Banerjee et al. 2013). It is worth to emphasize the importance of the poly(A) tail length, since it can affect the enzymatic degradation and the translationally dormant state of the mRNA in certain developmental stages (Jalkanen et al. 2014; Guhaniyogi & Brewer 2001).

The described 3' processing mechanism displays a high level of complexity that is fundamental to ensure its precise regulation, crosstalk between other steps of gene expression and also to define the proper definition of the cleavage site (Yang & Doublié 2011; Tian & Manley 2013; Moore & Proudfoot 2009). Furthermore, the correct poly(A) site is recognized through a tripartite mechanism coordinated by coupled interactions between 3' processing factors and RNA sequence elements (Millevoi & Vagner 2010; Yang & Doublié 2011; Shi 2012). These sequence elements are called *cis*-elements, and almost all the eukaryotic poly(A) signals contain multiple *cis*-elements that include the consensus sequence AAUAAA or variants of these, a U- or GU-rich downstream element (DSE), and a U-rich upstream stimulatory element (USE) (Shi 2012) (Figure 1).

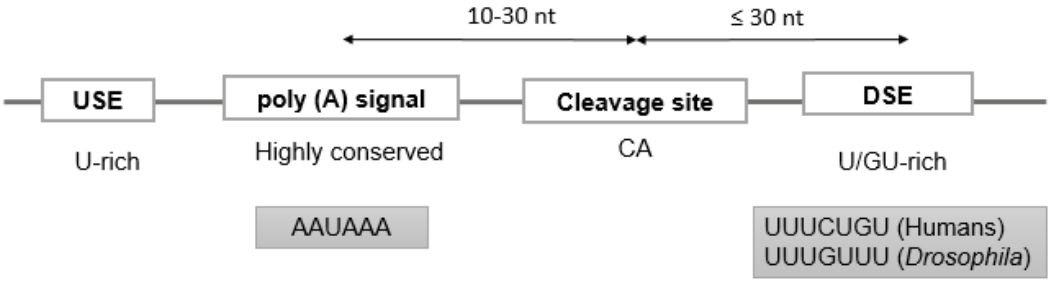

Figure 1 - Schematic representation of the cis-acting elements involved in 3'end formation

The poly(A) signal was first described as the hexamer AAUAAA by Proudfoot and Brownlee (Proudfoot & Brownlee 1976). Although this hexamer is present in the vast majority of the genes, variants also exist. In humans, the canonical poly(A) signal is present in 48-69 % of the genes and the most common variation of this signal is a single A-U nucleotide substitution on the second position (AUUAAA), which is present in 12-16% of the genes. Additionally, other variants are also present in the genome: 20% of these display single nucleotide variants in the AAUAAA hexamer and 10% do not have a recognizable AAUAAA-like sequence (Tian et al. 2005).

It is noteworthy that variations of the canonical poly(A) signal result in significant changes in cleavage and polyadenylation activity. In fact, it is described that the most efficient hexamer is the canonical one and all other non-canonical signals display a strong efficiency reduction. Interestingly, the only variant with a similar activity to the canonical poly(A) signal described is the AUUAAA (Wilusz et al. 1989; Sheets et al. 1990) (Figure 2).

Notably, a correlation between the type of poly(A) signal and its location can be made. In fact, it has been shown that the 3´ end of the transcripts display the highest frequency of AAUAAA hexamer (Hu et al. 2005). This difference in the sequence of the poly(A) signals supports the notion that the 3' most poly(A) signals are typically strong, ensuring proper termination of transcription, and upstream signals are weak, allowing regulation to occur.

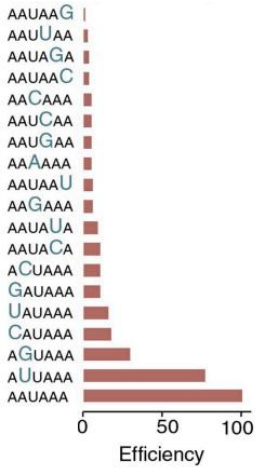

Figure 2- Diverse poly(A) signals and their polyadenylation efficiency (Adapted from Weaver 2012)

2.1. Alternative polyadenylation and gene expression

Most eukaryotic genes display mRNAs with multiple alternative 3' ends formed by polyadenylation at distinct sites, a phenomenon known as alternative polyadenylation (reviewed by Tian & Manley 2013; Lutz & Moreira 2010; Hsin & Manley 2012). In recent years it has become increasingly evident that alternative polyadenylation is much more pervasive than previously appreciated, since the latest estimate is that ~80% of human genes (Derti et al. 2012) and ~54% of *Drosophila* genes (Smibert et al. 2012) encode multiple transcripts derived from alternative polyadenylation.

Here is presented the classification of alternative polyadenylation in higher eukaryotes into three general events suggested by Lutz in 2008 (Figure 3). The *type I polyadenylation* involves the presence of one polyadenylation signal in the 3' UTR leading to only one resulting mRNA. In *type II alternative polyadenylation*, more than one polyadenylation signal is present in the common terminal exon, which leads to more than one mRNA isoform. In this type of polyadenylation the encoded protein is the same, however the expression may be affected by regulatory sequences present in the 3'UTR and cause changes in mRNA stability, localization or translation competence. The *type III alternative polyadenylation* refers to alternative

polyadenylation signals present in upstream introns or exons. In this way, the alternative polyadenylation is coupled with alternative splicing and it is possible to define type IIIi or type IIIe alternative polyadenylation regarding the position of the alternative polyadenylation signals, since they can be intronic (type IIIi) or exonic (type IIIe). These types of alternative polyadenylation may result in different protein products and also influence mRNA stability and translatability, or even lead to the insertion of an in-frame stop codon (Lutz 2008; Lutz & Moreira 2010).

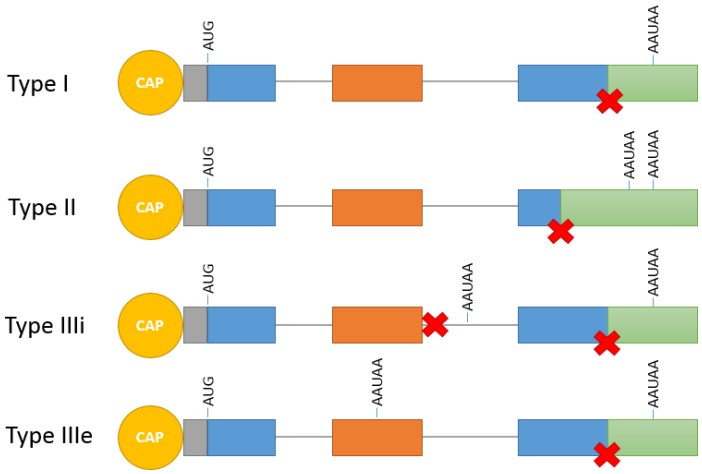

Figure 3- Schematic representation of polyadenylation events. Grey boxes, 5' untranslated regions; Green boxes, 3' untranslated regions; orange and blue boxes, coding regions; lines, introns.

Regarding the types of polyadenylation depicted, it is noteworthy that alternative polyadenylation contributes to increase the transcriptome complexity by producing isoforms with different coding sequence or 3′UTRs. Differential processing at multiple poly(A) sites in the 3'UTR could display important biological functions, since it plays a crucial role in gene expression's control (Fabian et al. 2010; Andreassi & Riccio 2009).

Alternative polyadenylation produces mRNA isoforms with different 3'UTRs containing different *cis*-acting elements. Since these *cis*-elements may harbour microRNA (miRNA)-target sites or binding sites for RNA-binding proteins the diversity of isoforms produced may affect gene expression quantitatively (Lutz & Moreira 2010). Therefore, longer 3' UTRs will more likely possess such elements and the mRNA will be more likely susceptible to regulation. Generally, shorter transcripts with shorter 3' UTRs produce higher levels of protein, which indicates that the amount of protein generated is dependent on the 3'UTR length (Mayr & Bartel 2009; Sandberg et al. 2008; Ji et al. 2011).

Another mechanism by which alternative polyadenylation affects protein expression is by regulating mRNA localization. In fact, alternative polyadenylation is able to modulate *cis*-elements found within the 3' UTR that dictated the localization of the mRNA (Andreassi & Riccio 2009). The control of the mRNA localization can be extremely important, for instance to promote rapid local protein synthesis in differentiated neurons (Di Giammartino et al. 2011).

The role of alternative polyadenylation in the regulation of different cellular states and programs such as proliferation, differentiation, and development programs is crucial, since it allows accurate regulation of a number of genes in a very precise manner (Lutz & Moreira 2010). In fact, through genome-wide analysis of alternative polyadenylation, it was possible to define a pattern that relates the proliferation and differentiation status of cells with the length of 3' UTRs. It was showed that proliferation and dedifferentiation are associated with a general shortening in 3' UTR length due to proximal poly(A) site usage, while 3' UTRs tend to be longer by distal poly(A) site usage during late developmental stages and differentiation processes (Rehfeld et al. 2013; Elkon et al. 2013).

Almost all the biological processes rely on accurate gene regulation in a precise temporal and spatial manner. As mentioned before, the alternative polyadenylation is a key player in the control of gene expression, therefore dysfunctional polyadenylation may be lead to a number of diseases and their progression (Mayr & Bartel 2009). Indeed, loss or gain of poly(A) sites in functionally important genes can be accomplished by genetic mutations or single-nucleotide-polymorphisms, and this phenomenon can cause or contribute to various diseases (Shi 2012; Curinha et al. 2014).

The association between diverse biological and pathophysiological processes and extensive alternative polyadenylation modulation indicates that poly(A) site selection is under precise and vigorous control. The regulatory mechanisms that control alternative polyadenylation are now coming into light and they form the rules that guide the poly(A) site choice that were coined by Yongsheng Shi as the "polyadenylation code" (Shi 2012).

2.2. Regulatory mechanisms of alternative polyadenylation

One well established alternative polyadenylation regulatory theme is the expression level of components of the 3′-end-processing machinery, since the levels of the core 3' processing factors can modulate the poly(A) site selection (Shi 2012; Elkon et al. 2013). A well-known example of this model of action occurs during B cell differentiation. It was described that during B-cell activation high protein levels of CstF64 (a core 3' processing factor) promote the usage of a proximal poly(A) signal in the IgM mRNAs that contains a weak CstF64 binding site. However, in resting B cells, which have a limited resource of CstF64, the distal poly(A) signal with a strong CstF64 binding site is preferentially recognized (Takagaki et al. 1996). Consequently, higher levels of 3' processing factors in proliferating cells may promote the recognition of weaker poly(A) sites, while the limited amounts of 3' processing factors in differentiated cells may be preferentially recruited to the stronger poly(A) sites (Shi 2012).

Several studies suggest that specific RNA-binding proteins (RBPs) are also involved in the control of alternative polyadenylation since they bind in close proximity to poly(A) sites, modifying the efficiency of cleavage/polyadenylation at those sites. One recent example is PABPN1 that suppresses weak poly(A) sites that results in an enhancement of polyadenylation

at stronger sites; since generally stronger sites are more distal, PABPN1 represses the proximal sites cleavage (Rehfeld et al. 2013; Shi 2012). Interestingly, a mutation of PABPN1 gene can cause the disease oculopharyngeal muscular dystrophy (OPMD), in which the most prevalent form of the disease results from a GCN expansion within in the first exon of the PABPN1 gene (Brais et al. 1998; Banerjee et al. 2013). Therefore, OPMD and PABPN1 provide an example of the development of a human pathology by a mutation in a core polyadenylation factor.

Alternative polyadenylation is likely to be also modulated by cell signalling pathways. Although little is so far known about this mechanism, a potentially interesting example is the process that up-regulates the levels of the protease thrombin under conditions of stress, which is achieved through 3' end processing regulation (Poort et al. 1996; Gehring et al. 2001; Danckwardt et al. 2011).

It is now appreciated that all the steps of mRNA biogenesis are intertwined and interconnected. The functional coupling between transcription and 3′ end processing is well documented and an explanation for this coupling is that it serves to increase the efficiency by which nascent transcripts are cleaved. However, the impact of this link on alternative polyadenylation regulation is just beginning to be elucidated. So far, some studies have emerged in regarding this association of transcription and alternative polyadenylation. In 2011, using a heterologous system, Nagaike *et al.* showed that transcriptional activators enhance processing efficiency and suggested that increasing the efficiency of 3' processing along transcribed genes would tend to favour the use of proximal poly(A) sites. In support of this model, it was also shown that the knockdown of the transcription elongation complex PAF1C lead to an accumulation of 3' extended transcripts of the target gene, since it resulted in the decrease of the 3′ cleavage and nuclear export of mRNA (Nagaike et al. 2011) .

At the same time, a second principle emerged regarding the interplay between transcription elongation rate and alternative polyadenylation with the work of Pinto *et al.* In this study, it was shown that Pol II elongation rate has an important role in poly(A) site selection, since a *Drosophila melanogaster* mutant strain, containing a point mutation in RbpI that results in a reduced transcription elongation rate, displayed an increase in proximal poly(A) site usage in a number of transcripts. A mechanistic explanation for these findings relies on the fact that proximal poly(A) sites are transcribed first and are encountered first by the 3′-end- processing machinery. Therefore, the use of proximal poly(A) sites should negatively correlate with transcription elongation rate (Pinto et al. 2011).

Interestingly, this scenario resembles to the effect that a lower transcriptional rate has on alternative splicing (de la Mata et al. 2003). Additionally, the work of Pinto *et al* clearly demonstrated a link between alternative polyadenylation and specific cellular states, since it assessed in *Drosophila* the physiological function of alternative polyadenylation in the 3'UTR at the organismal level (Pinto et al. 2011). Transgenic flies without the *polo* proximal (pA1) or distal (pA2) poly(A) signals were made and it was shown that flies that lack pA2 cannot express

sufficient levels of Polo protein and die in the pupa stage of development with severe abdominal abnormalities. The biological impact of pA2 deletion revealed that pA2 selection by the polyadenylation machinery is essential for abdominal histoblast proliferation, development of the adult epidermis and viability of the transgenic fly. Overall, this study highlights that *polo* alternative polyadenylation controls Polo protein levels and is tightly connected with cell cycle control, proliferation and the onset of metamorphosis (Moreira 2011).

It is well established the existence of multiple associations between the splicing and 3′-end-processing machineries. In the interplay between splicing and alternative polyadenylation regulation, it was demonstrated that splicing and polyadenylation factors interaction that occurs in terminal introns of pre-mRNAs enhance cleavage efficiency at 3′UTR poly(A) sites (Boelens et al. 1993; Ashe et al. 1995; Berg et al. 2012). Importantly, it has been suggested that there is a dynamic competition between splicing and polyadenylation, since the intronic polyadenylation events are usually associated with long introns with weak 5' splice sites, which require more time to be spliced out (Rehfeld et al. 2013). Another indication for the interaction between splicing and alternative polyadenylation regulation is that various splice factors have been shown to affect polyadenylation, such as NOVA (Licatalosi et al. 2008), PTB (Castelo-branco et al. 2004; Moreira et al. 1995) and U1 (Gunderson et al. 1994; Lou et al. 1998; Berg et al. 2012).

Epigenetic modifications of DNA structure have recently been implicated as having multiple impacts on RNA transcription and processing. In fact, it was described in the imprinted mouse gene H13 an allele-specific poly(A) site usage. This event is influenced by methylation primarily at CpG islands, since alleles without methylation of the CpG island utilize a proximal poly(A) site, generating a truncated H13, while alleles with methylation of the CpG island, utilize downstream poly(A) sites (Wood et al. 2008) .

Another factor implicated in alternative polyadenylation is nucleosome composition of the chromatin nearby the polyadenylation signals. It is described a decreased nucleosome density around poly(A) sites and an increased nucleosome density downstream of actively used poly(A) sites. Indeed, in genes with multiple poly(A) signals, a higher downstream nucleosome affinity was associated with higher polyadenylation signal usage (Rehfeld et al. 2013; Elkon et al. 2013).

All of the described regulatory mechanisms seem to be relevant and some authors suggest that poly(A) site selection is likely to be under the combinatorial control of multiple mechanisms. Therefore, currently the major challenge in deciphering the ‘‘polyadenylation code’’ is to integrate all of the different mechanisms in a quantitative manner. The importance of understanding such a code is highlighted by the impact of alternative polyadenylation in gene expression in health and disease.

3. The interplay between transcriptional elongation and alternative polyadenylation regulation

Alternative polyadenylation is a well-orchestrated process that presents plasticity, while still displays high specificity and fidelity. Therefore, alternative polyadenylation is an event with an accurate, multilayer and complex regulation, where transcription is one of the players (Danckwardt et al. 2008; Dujardin et al. 2013).

Alternative polyadenylation and transcriptional elongation are spatiotemporally coordinated and even though great effort has been made to characterize the interplay between both events most of the questions still remain unanswered. Nevertheless, some insight on this subject have been achieved by applying the knowledge already obtained on the coupling between alternative splicing and transcription, since some valid and similar conclusions can also be applied to alternative polyadenylation (Moreira 2011).

It is well established that alternative splicing is regulated by the relative abundance of splicing factors and also by a more complex process involving the transcription machinery. Indeed, transcription and splicing can occur co-transcriptionally and in certain situations splicing is tightly coupled to transcription (Kornblihtt et al. 2013). In order to explain how this coupling works two non-mutually exclusive models have been proposed based on specific recruitment of Pol II-associated factors, the recruitment coupling model, and on splicing regulation by transcription elongation rates, the kinetic coupling model (Nogués et al. 2003). Excitingly, growing evidence suggest that the named models might also apply for coupling of alternative polyadenylation and transcription.

3.1. The kinetic model

The kinetic coupling model predicts that the rate of Pol II elongation influences the outcome of splicing. This model was supported by numerous experiments from the Kornblihtt group, namely de la Mata's work showing that decreased Pol II elongation rate stimulates fibronectin EDI inclusion (~4 fold) (de la Mata et al. 2003). Therefore, the kinetic model suggests that a slower elongation rate favour the inclusion of alternative exons governed by an exon skipping mechanism, while a fast elongation rate favours exclusion of these. Additionally, many studies showed that kinetic coupling depends on the sequence context and the specific combination of splicing regulators that are involved on the accessibility of *cis*-competing splice sites in pre-mRNA (de la Mata et al. 2003; Nogués et al. 2003). Truly, the influence that elongation rates have on splicing can be interpreted as being consistent with the 'first come, first served' model, first postulated 26 years ago (Aebi & Weissman 1987).

In 2011, Pinto *et al.* showed that in addition to affect alternative splicing, the Pol II elongation rate also plays a role in alternative polyadenylation (Pinto et al. 2011). In this study,

the direct function of Pol II elongation rate in *polo* poly(A) signal selection was studied in *Drosophila melanogaster* using the C4 fly mutant, as this shows a 50% slower transcription elongation rate than its wild-type counterpart. The *polo* gene is involved in many critical steps in the cell cycle (Sunkel & Glover 1988) and generates two mRNAs that differ in their 3'UTR length due to the usage of two tandem pA signals, AUUAAA (pA1, proximal) and AAUAUA (pA2, distal). Therefore, the two *polo* transcripts have the same open reading frame and code for the same protein (Llamazares et al. 1991).

By using the C4 fly mutant the authors were able to show that when the elongation rate is reduced by 50%, the proximal poly(A) signal usage is increased by 3.5 fold than in the wild type. Presumably, transcription by the "slow" Pol II allows longer exposure of the pA1 signal in the nascent pre-mRNA that will be thus recognized by the polyadenylation machinery before Pol II reaches the distal pA signal (Figure 4). The influence that elongation rates have on alternative polyadenylation can be interpreted as being consistent with the 'first come, first served' model, first postulated for alternative splicing (Aebi & Weissman 1987). Moreover, using this mutant Pol II, a comparison can be drawn between the value obtained for the increase of proximal poly(A) signal usage (3.5 fold) (Pinto et al. 2011) and for alternative EDI exon inclusion (4-fold) (de la Mata et al. 2003).

Indubitably, the most direct support for kinetic coupling has come from the use of C4 fly mutant , which harbour an amino acid substitution (change in amino acid 741 from Arg to His) in the *RpII215* gene that encodes the largest subunit of Pol II and show a reduced elongation rate (Coulter & Greenleaf 1985; Chen et al. 1996). After the C4 fly mutant contribution to elucidate the kinetic interaction between alternative polyadenylation and elongation, a pertinent question raised: "Is the alternative polyadenylation pattern affected by an enhancement of Pol II elongation rate?". Even though Pinto *et al*. suggested that a "fast" Pol II would favour the distal poly(A) signal usage (Figure 4), this is still an enigma. Interestingly, a key tool to address this question would be a fly mutant able to display a phenotype opposite to that presented by C4. That mutant is the S1 fly mutant, which carries a mutation that alters amino acid 728 from Ser to Cys in *RpII140* gene which encodes the second largest subunit of Pol II and display a ~50% enhanced Pol II elongation rate (Coulter & Greenleaf 1985; Chen et al. 1996). Truly, with the tools to answer this question, it can only be imagined the new challenges that will appear when exploring this flied.

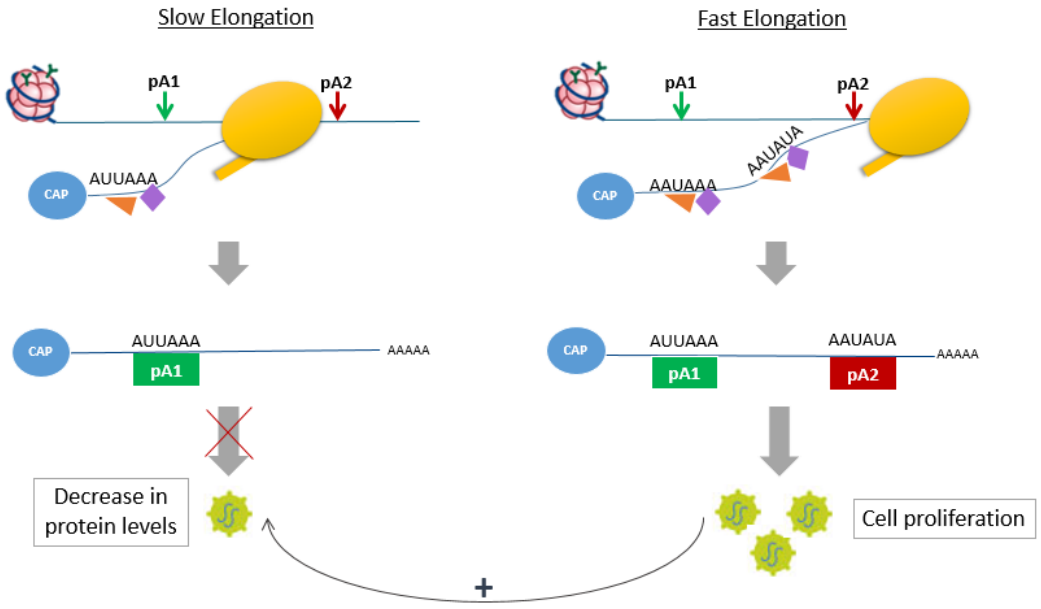

Figure 4- Proposed model by Pinto et al for the role of Pol II kinetics on *polo* poly(A) site selection (Adapted from Pinto et al 2011 and Moreira 2011)

3.2. The recruitment model

The recruitment coupling model implies that the recruitment of factors by the transcription machinery will affect the subsequent splicing decisions. Examples of this model are transcription elongation regulator 1 (TCERG1) and DBC1–ZIRD (DBIRD) complex that are able to modulate selective exon skipping through the control of Pol II elongation (Montes et al. 2012; Close et al. 2012).

In comparison to alternative splicing events, it is also possible that the coupling between alternative polyadenylation and transcription elongation occurs through recruitment of specific transcription elongation factors that influence the poly(A) site usage. Accumulating evidence indicates that transcriptional activity impacts poly(a) signal choice, since defects in the transcriptional elongation factors RPB2, TFIIS, and SPT5 enhance usage of upstream poly(a) site in yeast (Cui & Denis 2003) . Additionally, ELL2 (eleven-nine- teen lysine-rich leukemia gene 2) has been shown to modulate the usage of the proximal poly(A) site and increase exon skipping in IgH, by competition between the mechanisms of alternative polyadenylation and splicing (Martincic et al. 2009).

The Martincic *et al*. and Cui and Denis's work suggest an impact of elongation factors in poly(A) site choice and revealed an all new world of crosstalks between these events. Consequently, there is a growing need for studies to assess the role of other transcription elongation factors in alternative polyadenylation in other models, like *Drosophila,* using more suitable genes (for instance, genes without splicing - e.g. the *polo* gene). In this thesis, it is proposed to study the impact of both elongation factors dELL and dSpt6 in *polo* poly(A) site

selection, since they both display distinct and crucial roles in transcription elongation and therefore would allow a new insight in the mentioned interplay.

a) dELL

Originally isolated from rat liver nuclear extracts, ELL was first identified as a fusion partner of the MLL (mixed lineage leukemia) gene in chromosomal translocations associated with mixed lineage leukaemias in children and shown later to increase elongation by reducing the rate of Pol II stalling during active transcription in acute myeloid leukemia (Thirman et al. 1994; Shilatifard et al. 1996).

To this date, three ELL family members (ELL1, ELL2 and ELL3) have been described in mammalian cells, all with biochemical elongation activity but differing in expression pattern (Thirman et al. 1994; Shilatifard et al. 1996). In *Drosophila melanogaster* just a single ELL homolog has been identified (dELL) and although it displays significant sequence homology to all three mammalian ELL family proteins it is most closely related with ELL2. Its ubiquitous expression at all developmental stages suggests that dELL has a widespread role in gene transcription during development (Gerber et al. 2001).

In addition to its structural homology to mammalian ELL family proteins, dELL behaves like a functional homolog of mammalian ELL, since it is a nuclear protein that can increase the elongation rate of Pol II and it physically interacts with this enzyme (Gerber et al. 2001; Byun et al. 2012). This elongation factor has been found to participate in distinct complexes as SEC and Little Elongation Complex (LEC) and also proved to have nonredundant functions *in vivo* (Smith, Lin & Shilatifard 2011; Smith, Lin, Garrett, et al. 2011) .

The impact of dELL in alternative polyadenylation is therefore urgent to investigate due to its function in the enhancement of Pol II elongation rate and its role revealed by the Martincic *et al*.'s work that highlight the connection between the mammalian homolog ELL2 with polyadenylation and exon skipping (Martincic et al. 2009).

b) dSpt6

Another important molecule in the interplay between alternative polyadenylation and elongation is Spt6 (Suppressor of Ty 6), which was one of the first known elongation factors for which the role in elongation rate was robustly confirmed *in vivo* (Ardehali et al. 2009).

The *spt6* gene was initially identified from the *Spt* genetic screen in yeast. This elongation factor has been shown to be present on the body of genes in various model systems in a transcription-dependent manner (Saunders et al. 2006; Kim et al. 2004; Kaplan et al. 2005). Further evidence for the *Drosophila* Spt6 (dSpt6) role in transcription elongation showed that it is a *bona fide* transcription elongation factor, which positively stimulates the elongation rate of Pol II *in vivo*, co-localizes with the Ser2 phosphorylated form of CTD on polytene chromosomes, and

it is critical for normal development and morphogenesis throughout the life cycle of *Drosophila* (Andrulis et al. 2002; Yoh et al. 2007; Ardehali et al. 2009; Dronamraju & Strahl 2014). In addition, dSpt6 can regulate histone modifications by acting as a platform for histone modifiers or by performing a molecular connection between histones and template DNA (Kato et al. 2013). Taking previous studies into account, Spt6 is also suggested to be critical to transcription termination (Ardehali et al. 2009).

The interplay between transcription elongation and alternative polyadenylation ought to be quite complex. Therefore, a complete understanding of the role of elongation factors, such as Spt6, and affect poly(a) signal selection is crucial to expose such elusive connection.

3.3. The integration between kinetic and recruitment models

The coexistence of both models (kinetic and recruitment) has been more recently proposed in an integrated model for splicing. Indeed, some findings suggests a strong connection between transcription and splicing that might be the consequence of a combination of the two models discussed above, since some splicing factors are able to alter the transcriptional elongation rate and introns are necessary for efficient Pol II transcription (Nogués et al. 2003).

Truly, the interplay between alternative polyadenylation and elongation is likely to also be integrated both by elongation rates and elongation factors, in a bidirectional crosstalk between the kinetic and recruitment models (Moreira 2011). The interdependence between elongation and alternative polyadenylation events is emphasized by the finding that mutations in cleavage/polyadenylation factors lead to inefficient transcriptional elongation and processivity (Luna et al. 2005).

It is well known that the Pol II enzyme exhibits a dynamic behaviour by displaying different rates, CTD modifications, differential recruitment of accessory factors or even by undergoing reversible modifications. Interestingly, the present panorama of insight suggests that this dynamic behaviour of Pol II can affect alternative polyadenylation regulation in a bidirectional-manner. Definitely, it is the interplay between the dynamic behavior of Pol II and alternative polyadenylation that holds one of the most interesting questions in the field: how mechanistically alternative poly(A) sites are selected.

Chapter 2

Objectives

The work described in this thesis was aimed to understand the interaction between transcriptional elongation and alternative polyadenylation. This was addressed using the *Drosophila polo* as a gene model, by assessing the impact of a mutant Pol II with an increased elongation rate and by depletion of the elongation factors dELL and dSpt6 in *polo* poly(A) site choice.

- This page was intentionally left blank. –

Chapter 3

Material and Methods

1. Drosophila melanogaster stocks

The fly stocks w^{1118} (wild-type), *RpII215* and *RpII140* were obtained from the Bloomington Stock Centre. The *RpII215* flies carry the C4 mutation in RNA polymerase II 215kD subunit that modifies the amino acid at position 741 from Arg to His (Chen et al. 1993). The *RpII140* flies carry the S1 mutation in the RNA polymerase II 140kD subunit, modifying the amino acid at position 728 from Ser to Cys (Kim et al. 1994).

All fly stocks were grown at 25ºC using standard cornmeal/honey/agar yeast medium and maintained by periodically transferring adults to vials with fresh medium.

1.1. *Drosophila* Schneider (S2) cells

The *Drosophila* Schneider (S2) cells were first isolated in 1972 from a primary culture of late-stage *Drosophila melanogaster* embryos, and many features of this cell line suggest that it is derived from a macrophage-like lineage (Schneider 1972; Zanoni et al. 2009). Today, the relevance/potential of this cell line in the molecular biology field is undeniable.

a) Cell culture

Cells were grown as a loose, semi-adherent monolayer in 25 cm^2 tissue culture flasks (Sarstedt) or 6 well plates (Frilabo) and maintained at 25ºC without CO_2 in Schneider's Insect Medium (with L-glutamine and sodium bicarbonate, Invitrogen) supplemented with 10% (v/v) Fetal bovine serum (FBS) (Invitrogen).

b) Cryopreservation of S2 cells

Cells were grown to sub-confluency (approximately 1-2 x 10^7 cells/ml in 5 ml), and harvested by centrifugation at 1200 rpm for 5 min. Regular growth medium was discarded and cells were resuspended at approximately 1.1 x 10^7 cells/ml in freezing medium (Schneider's Insect Medium with 10% FBS and 10% Dimethyl sulfoxide (DMSO)). The cell suspension was aliquoted to cryovials (Simport) and stored first at -20°C for 24h and finally at -80°C for long term storage.

c) Thawing of S2 cells

The S2 cells cryovial was quickly placed in a 37°C water bath and the cells were transferred to 5 ml of warm regular growth medium immediately after thawing. The cells were centrifuged at 1200 rpm for 5 min, resuspended in 5ml of regular growth medium and finally transferred to a 25 cm^2 cell culture flask.

2. Total RNA extraction

Total RNA isolation from flies or S2 cells was performed with TRIzol Reagent (Invitrogen) according to manufacturer's instruction with some modifications. Briefly, 20 male adult flies were collected and frozen in 200 µL of TRIzol at -80°C. After thawing, flies were homogenized with a disposable plastic pestle and 800 µL of TRIzol was added. The isolated RNA was quantified in a Nanodrop 1000 Spectrophotometer (Thermo Scientific) by measuring the absorbance at 260 nm (A_{260}).

In order to eliminate any possible DNA contamination, 1 µg of the newly isolated RNA was treated with treated with Deoxyribonuclease I (DNAse I) (Roche) for 30 minutes at 37°C, according to manufacturer's instruction. The reaction was stopped by adding [[2-(Bis-carboxymethyl-amino)-ethyl]-carboxymethyl-amino] acetic acid (EDTA, pH 8.0, final concentration 8 mM) and incubating for 10 minutes at 75°C, to heat inactivate the DNase I. Additionally, the integrity of the RNA isolated was assessed using Experion (BioRad).

2.1. cDNA synthesis

Complementary DNA (cDNA) was synthesized using the 1 µg of DNAse I treated RNA by adding 1 µl of Random Hexamers (50 µM, Sigma-Aldrich), 1 µl dNTPs (10 mM, Thermo Scientific) and RNAse free water (Thermo Scientific) to a final volume of 14 µl. The mixture was incubated for 5 minutes at 65°C and 4 µl of 5x cDNA synthesis buffer (Invitrogen), 1 µl of Dithiothreitol (DTT, Invitrogen), 0.5 µl of RiboLock RNase Inhibitor (40 U/µL, Thermo Scientific)

and 0.5 µl of SuperScript III Reverse Transcriptase (200 U/µL, Invitorgen) was added to a final volume of 20 µl. Afterwards, the mixture was incubated for 5 min at 25°C, 60 min at 50°C and 15 min at 70°C to in a 48-well TPersonal Thermocycler (Biometra). cDNA was stored at -20°C.

2.2. Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) was performed by adding 1 µg of cDNA, 2 µl of $MgCl_2$ (25 mM), 1 µl dNTPs (10 mM), 1 µl of each oligonucleotide (10µM) (Table A1, see Appendix), 2 µl of Green GoTaq Reaction Buffer (Promega), 2 µl of GoTaq DNA Polymerase (5 U/µl, Promega) and double-distilled sterile water (ddH2O) up to 20 µl. The reaction was performed in a 48-well TPersonal Thermocycler (Biometra), with the thermal cycling conditions of 95°C for 2 minutes, followed by 30 cycles of: 1 min at 95°C, 30 sec at the optimized temperature for each primer pair and 72°C for 1 min. A final extension was performed at 72°C for 5 min. Furthermore, PCR products were analysed by gel electrophoresis on 1.5% Tris-acetate-EDTA (TAE) agarose gel stained with SYBR safe (Life Technologies).

2.3. Real-time PCR reaction (qPCR)

Gene expression quantifications were performed by real-time PCR reaction (qPCR). Each reaction contained 1 µL of the synthesized cDNA, 5 µL SYBR Select Master Mix (Life Technologies) and ddH2O up to 10 µL, with variable concentrations of each oligonucleotide pair: 0.15 µM of *polo total*, *CG6024 total* and *7SL*; and 0.5 µM of *polo pA2* and *CG6024 pA2*.

The primers *polo total* and *CG6024 total* were designed to anneal in the end of the coding region, which allowed to measure all the mRNA isoforms produced by each gene. Additionally, these primers anneal to exons on both sides of an intron to allow differentiation between amplification of cDNA and potential genomic contamination. The *polo pA2* and *CG6024 pA2* primers were designed to allow only the quantification of the long isoform of each gene.

The samples were run on a StepOne Real-time PCR System (Applied Biosystems) using the following program: 50ºC for 2 min, 95°C for 2 min, 40 cycles of 95°C for 15 sec and 1 min 58°C. The instrument was also programmed to include a melting profile immediately following the thermal cycling protocol, in order to create a complete melting curve.

It should be highlighted that cDNAs were tested in triplicate and “no template” controls were always analysed to ensure that the signal generated was specific and not derived from cDNA only and not from primer-dimers, genomic DNA, or any of the cDNA reagents.

The qPCR reaction efficiency using specific oligonucleotides was evaluated by performing a standard curve with 10-fold serial dilutions of cDNA. In an optimal reaction the DNA double in each cycle (n) and consequently the dilution factor is equal 2^n (Bio-Rad Laboratories 2006). Thus, with a perfect efficiency (E=2) the 10-fold serial dilutions standard curve’s slope is -3.32, since $2^{3.32} = 10$. Equation 1 was used to calculate the qPCR efficiency using this primer pair

(E) and only when the efficiency is between 1.9 and 2.1 the reaction is considered efficient (Bio-Rad Laboratories 2006).

$$E = 10^{(-1/slope)} \qquad \textit{Equation 1}$$

Generally, two quantification strategies are applied to gene expression quantification via qPCR, an absolute or a relative quantification (Pfaffl 2004), the latter was used in this study.

The relative expression of a gene can be calculated with Equation 2:

$$\text{Relative expression} = 2^{-\Delta Ct} = 2^{-[Ct\ \text{target gene} - Ct\ \text{reference gene}]_{\text{test condition}}} \qquad \textit{Equation 2}$$

Where C_t is the number of cycles required to reach a fluorescence threshold during the exponential phase of amplification (Livak & Schmittgen 2001) . The $-\Delta C_t$ method provides the expression of a target gene in relation to a reference gene and assumes a maximal reaction efficiency (100%, E = 2).

Additionally, mathematical equations based on the comparison of the distinct cycle differences are also established to assess the relative expression ratio (R) (Equation 3). If maximum primer efficiency is assumed ($E_{target}=E_{reference}=2$), then the Equation 3 can be presented in a simpler form (Equation 4). Additionally, under the conditions of Equation 5, the Equation 4 can be simplified in Equation 6 where the amount of target is normalized to an endogenous reference and relative to a calibrator (non-treated or scramble transfected cells). The explained process is the $-\Delta\Delta C_t$ method and it generates a normalized fold expression, in which values higher than 1 mean an increase of expression of the target gene in the specific condition relatively to the calibrator, while values lower than 1 translate decrease of expression.

$$R = \frac{(E\ target)^{\Delta Ct\ target\ (control-sample)}}{(E\ reference)^{\Delta Ct\ reference\ (control-sample)}} \qquad \textit{Equation 3}$$

$$R = \frac{2^{-[Ct\ test\ gene - Ct\ reference\ gene]\ test\ condition}}{2^{-[Ct\ test\ gene - Ct\ reference\ gene]\ reference\ condition}} \qquad \textit{Equation 4}$$

$$-\Delta Ct = -[\Delta Ct\ sample - \Delta Ct\ control) \qquad \textit{Equation 5}$$

$$R = 2^{-\Delta\Delta Ct} \qquad \textit{Equation 6}$$

In the present work, the results were presented using a relative quantification approach. The data was analysed using the "$-\Delta C_t$ method" and "$-\Delta\Delta C_t$ method", assuming the maximum efficiency (E= 2).

3. RNA interference

RNA interference (RNAi) is a biological mechanism by which double-stranded RNA (dsRNA) induces gene silencing by targeting complementary mRNA for degradation (Montgomery 2004). Undoubtedly, this mechanism changed the way researchers study gene function and it is an extremely useful tool in the molecular biology field. Luckily, performing RNAi assays in S2 cells is simpler than in other cell lines, since these are phagocytic cells that can easily take up dsRNA from the cell culture medium (Rocha et al. 2011).

In this project, gene silencing by RNAi was performed in Drosophila S2 cells to assess the possible role of dELL or dSTP6 in *polo* polyadenylation signal selection process. The experimental procedure is depicted in Figure 5.

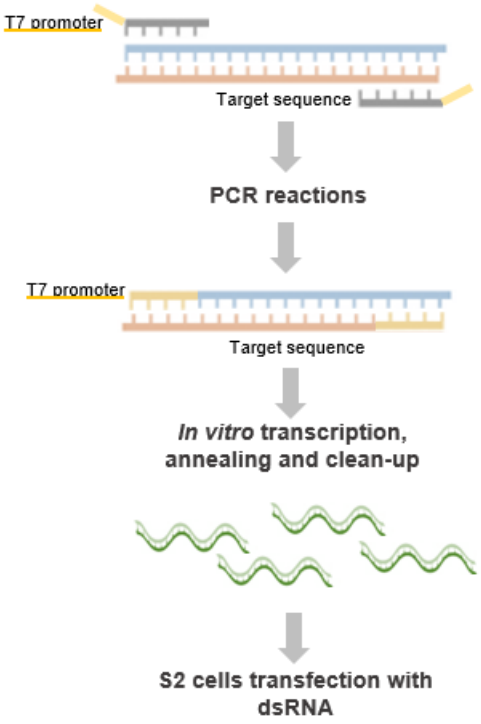

Figure 5- Schematic representation of RNAi experimental procedure

3.1. Preparation of DNA template

Double-stranded RNA was *in vitro* transcribed from DNA templates obtained from PCR products generated by oligonucleotides flanked with T7 RNA Polymerase promoter site on both ends. Table A2 (see Appendix) lists all the primers used with their respective sequences. PCR reactions were performed as described previously (Materials and Methods, section 5) to a final volume of 50 µL. The PCR programs used varied accordingly to each oligonucleotide pair requirements (Table A3, see Appendix).

The PCR products obtained were purified using the QIAquick PCR Purification Kit (Qiagen), accordingly with the manufacturer's instructions and examined on a 1.5% TAE-agarose gel.

3.2. Production of dsRNA

The design of the dsRNA molecules was carefully executed for the mRNA of each elongation factor studied (dELL or dSpt6), taking into account the possible unspecific targeting, targeting of all isoforms and the strongest expression reduction. In addition, a negative control was used (*DsRED*) which did not displayed homology to any of the target mRNAs in the cells.

In vitro transcriptions were performed using the Ambion Megascript kit (Life Technologies) with each reaction containing 1µg of DNA template, 10µl of NTPs (75mM), 2.5µl of 10x Reaction Buffer, 2.5 µl of 10x enzyme mix and RNAse free water (Thermo Scientific) up to 25 µL. The mixture was incubated overnight (16 hours) at 37°C, followed by a 15 min (37°C) incubation with 1 µl of turbo DNAseI (Ambion Megascript kit) and finally purified with the PureLink RNA Mini Kit (Life Technlogies). Annealing of the newly *in vitro* transcribed RNA was performed by a slow annealing program that started with an incubation of 5 min at 96°C and then sequentially decreased 2°C each minute until it reached 24°C. Sample where then diluted (1:10) in H_2O and 1 µl and 3 µl of the diluted material were examined on a 1.5% TAE-agarose gel.

3.3. Transformation of S2 cells

For these experiments exponentially growing S2 cells just about to reach confluence were used. Cell density was measured and calculated using a Countess Automated Cell Counter (Invitrogen) and 1×10^6 cells per ml of serum free medium was added to each well of a 6-well plate. The dsRNAs (Table 2) for each target gene were added and gently mixed by swirling the plate. The cells were allowed to settle for 1 h in the incubator. Afterwards, 2 ml of 10% FBS (Invitrogen)-containing Schneider's medium (Invitrogen) was added to each well. The plate was then further incubated at 25°C for different periods of time (Table 2).

Table 2- Amount and incubation period of each dsRNA used in the RNAi assays.

Target gene	**dsRNA amount**	**Incubation period**
dell	25µg	120 hours
dspt6	25µg	96 hours
DsRED	25µg	120 hours (as dELL's control)
		96 hours (as dSpt6's control)

4. Statistical Analysis

The significance of differences in gene expressing was determined by two tailed Student's t-test analysis with 95% confidence intervals. This was performed using GraphPad Prism version 6.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Differences were considered to be statistically significant if $p<0.05$.

- This page was intentionally left blank. –

Chapter 4

Results

1. Assessment of the impact of Pol II elongation in alternative polyadenylation

In the present work, it was examined the alternative polyadenylation pattern of p*olo* and *CG6024* genes in an enhanced Pol II elongation rate scenario, aiming to improve the current knowledge in this field and therefore complete the kinetic model proposed by Moreira's group (Pinto et al. 2011; Moreira 2011).

1.1. RNA extraction optimization

In order to determine the alternative polyadenylation pattern of different genes an initial RNA extraction from flies was performed. In this work, RNA extraction was executed as described in the previous section (see *Material and Methods*), however the DNAse treatment was initially done using all the extracted RNA and the incubation period was up to 1hour. In these conditions, the RNA obtained displayed high levels of degradation (Figure 6), which did not allow its use in further assays.

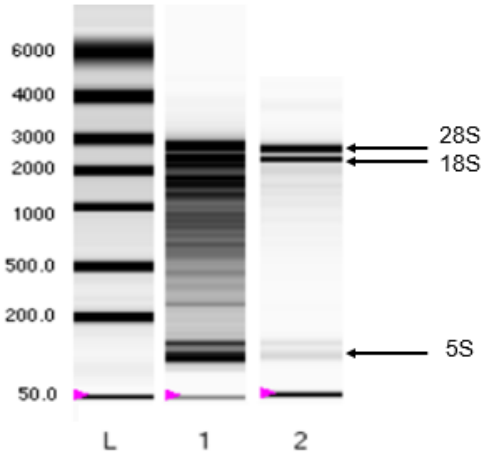

Figure 6- RNA quality assessed by Bio-Rad Experion electrophoresis system. **(L)** Molecular weight ladder **(1)** RNA extracted and treated with DNAse up to 1hour **(2)** Sample of high quality RNA.

In order to overcome the obstacle of RNA degradation, different RNA extraction methods were tested (Table 3) and different variables were studied: the DNAse enzyme, the DNAse incubation period, the DNAse heat inactivation (10 minutes incubation at 75ºC), the DNAse inactivation with EDTA (pH8, final concentration 8 mM) and the use of Ribolock (Thermo Scientific).

Table 3- Conditions used in the optimization of the RNA extraction process

Condition	Description
1	RNA sample with 15 min DNAse treatment and inactivation with heat and EDTA
2	RNA sample with 30 min DNAse treatment and inactivation with heat and EDTA
3	RNA sample with 60 min DNAse treatment and inactivation with heat and EDTA
4	RNA sample with 90 min DNAse treatment and inactivation with heat and EDTA
5	RNA sample with 60 min DNAse treatment and inactivation with only with heat
6	RNA sample with Ribolock, 60 min DNAse treatment and inactivation with heat and EDTA
7	Mammalian RNA sample with 60 min DNAse treatment and inactivation with heat and EDTA
8	RNA sample treated with Turbo DNAse (Life Technologies) (used according to manufacturer's guideline)
9	RNA sample without DNAse treatment

The RNA quality in each condition tested was assessed using the Experion system (Figure 7). This system allows rapid and automated results, it requires only a very small amount of RNA sample and the integrity of the RNA can be assessed by visualization of the 28S (two similar-sized bands in *Drosophila*) and 18S ribosomal RNA bands (Fleige & Pfaffl 2006; Winnebeck et al. 2010).

In Figure 7, it is possible to observe that the RNA in the sample 9 (not treated with DNase) is not degraded, which indicates that, before the DNase treatment, the RNA quality is acceptable and the treatment is critical for RNA degradation. Additionally, a different enzyme was used (Turbo DNAse (Life Technologies), kindly provided by Professor Jorge Vieira from IBMC; sample 8) to determine if the RNA degradation was due the enzyme used. Since some samples tested with the regular enzyme did not displayed any RNA degradation, it was excluded the hypothesis that the DNase could favour RNA degradation. Notably, between all the incubation periods tested, 30 minutes (sample 2) was selected as the most preferable one, since it displayed the lower level of degradation and did not showed any genomic contamination. In contrast, longer incubation periods (sample 3 and 4) exhibited high levels of degradation. Another variable tested, was the EDTA used to inhibit DNase after treatment. It was proved that the EDTA role is crucial, since samples without it (sample 5) display higher levels of degradation in comparison with samples in the same condition but treated with EDTA (sample 3). Additionally, it was showed that the protective role of Ribolock against RNases is not necessary, since sample 6 showed the same degree of degradation as the sample without Ribolock (sample 3). To exclude the possibility that the poor RNA quality was due the sample type or extraction method, a mammalian sample (sample 7) was treated in the same conditions as sample 3. Since the degradation levels of both samples was almost the same the named possibility can be ruled out.

These results allowed to successfully determine the condition in which the RNA extraction must be performed in order to achieve RNA that is not degraded and that does not display genomic contamination. In conclusion, 30 minutes of DNase incubation must be used, the enzyme should be inactivated with heat and EDTA and Ribolock is not necessary.

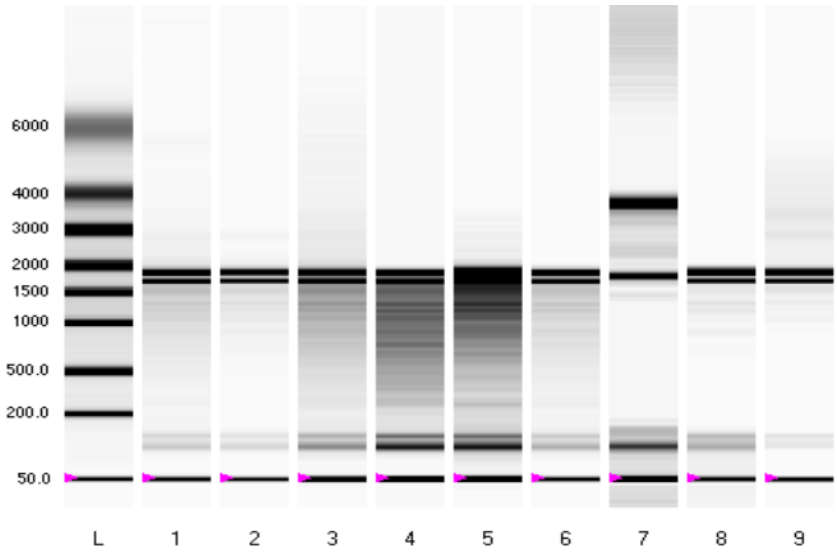

Figure 7- RNA integrity of different RNA extraction conditions tested. (L) Molecular weight ladder **(1) to (9)** RNA extracted and treated according to the conditions described in Table 3.

1.2. Assessment of alternative polyadenylation in polo and CG6024

Using RNA extracted from *w* ^{1118}, *RpII215* and *RpII140* flies and specific primer pairs, the role of Pol II kinetics on poly(A) site usage of *polo* and *CG6024 g*enes was studied by qPCR.

a) *polo* gene

In order to study the alternative polyadenylation pattern of *polo* in the different *Drosophila* strains, primers that differentially measure the total mRNA level and the long mRNA isoform were used (*polo total* and *polo pA2* primers) (Figure 9 a). The qPCR efficiency using each primer pair was assessed in different conditions, allowing to find the one in which the efficiency was between 1.9 and 2.1.

Regarding the *polo total* primer pair, it was previously shown that 58°C could be used as annealing temperature (Henriques 2011). Therefore, the first attempt to find the efficient qPCR conditions for *polo total* primer pair, was using 0.15 µM of each primer and 58ºC as annealing temperature. The efficiency in this condition was calculated as indicated by W. Paff (see *Materials and Methods*) and value obtained was 2.07 (Pfaffl 2001). Additionally, the standard curve used to assess qPCR's efficiency using *polo total* primers is display in Figure 8, as an example of the approach used.

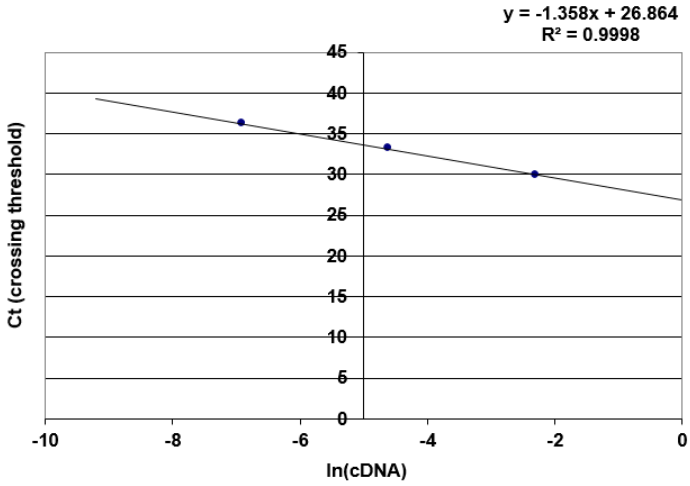

Figure 8- Standard curve for *polo total* primer pair. A plot of C_T against the ln of cDNA, results in a straight line. This is a linear regression generated using ten-fold serial dilutions of template and the qPCR efficiency using *polo* primer pair is determined from the slope of the calibration curve.

Concerning the *polo pA2* primer pair, it was also previously shown that 58°C could be used as annealing temperature (Henriques 2011), therefore this temperature was used and different primer concentrations were tested. Initially, the qPCR efficiency using *polo pA2* primer pair was assessed in the same condition stated to *polo total* primer, however the efficiency obtained was 2.33, thus out of the desirable range. Then, the primer concentration was raised to 0.5 µM, which lead to an efficiency of 1.95.

In addition to optimize the conditions of *polo total* and *polo pA2* primers, the efficient condition using *7SL* (reference gene) primers must be determined. In fact, the same conditions used to *polo total* primer pair were tested with *7SL* primers and the efficiency achieved was 2.0.

After measuring and optimizing the qPCR efficiency for each primer pair, qPCRs were performed to assess the levels of *polo* total and pA2 in *RpII215*, *w* ^{1118} and *RpII140* flies.

Figure 9 shows that the *polo* mRNA levels differ between the fly strains tested. The similar mRNA levels of total *polo* and the decrease in *polo* pA2 displayed by the *RpII215* mutant in comparison with *w* ^{1118} are consistent with previous work (Pinto et al. 2011). The novelty was the comparison between *polo* total mRNA levels of *RpII140* flies and the wild-type counterpart, since *polo* levels decreased 0.6 fold (i.e. 60%) in the flies with an enhanced Pol II elongation rate (*RpII140*). Regarding *polo* pA2 mRNA levels, the *RpII140* mutant showed similar levels to *w*^{1118} (1.08 fold, i.e. 108%). These results suggest that *polo* short mRNA isoform (*polo* pA1) is decreased in *RpII140*, since *polo* total levels decreased but *polo* pA2 levels were maintained. Noteworthy, in *RpII140* flies this outcome affects the *polo* proximal poly(A) site (pA1), which is the stronger one (AUUAAA).

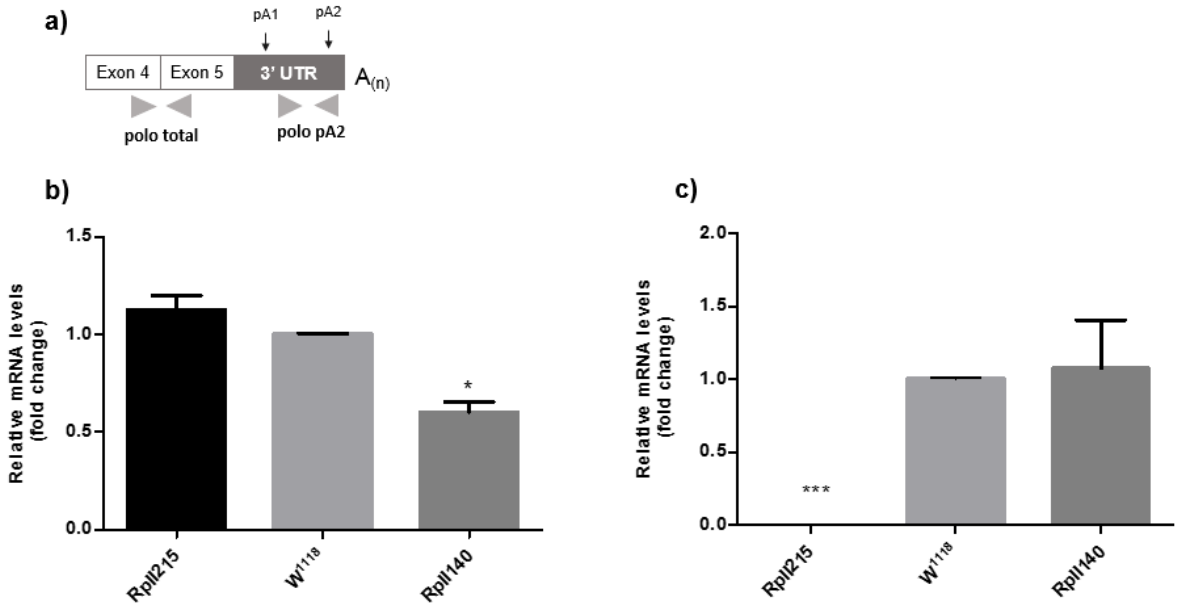

Figure 9-RNA Pol II kinetics with a slower an enhanced transcription elongation rate affects *polo* alternative polyadenylation. a) diagram that illustrate primer positions for qPCR analysis **b)** *polo* mRNA levels in *RpII215*, *w* ^{1118} and *RpII140* flies and a **c)** *polo* pA2 mRNA levels in *RpII215*, w ^{1118} and *RpII140* flies. Graph represents fold change of *polo* mRNAs in adult flies, relative to 7SL mRNA. For all the panels, error bars show standard error of the mean (s.e.m.) from at least three independent experiments

b) *CG6024* gene

The *CG6024* gene shows a similar genomic structure to *polo*, since it contain two poly(A) signals in the 3'UTR (information kindly provided by Joel Graber and GenBank, NCBI). Therefore, in order to study the alternative polyadenylation pattern of *CG6024*, primers able to differentially measure total polo mRNA level and the long mRNA isoform were used (*CG6024 total* and *CG6024 pA2* primers).

It was previously shown that 58ºC could be used as annealing temperature of both oligonucleotides *CG6024 total* and *CG6024 pA2* , therefore that was the temperature used in this work (Henriques 2011). The efficiency obtained for the *CG6024 total* primer pair, when tested in the same condition stated to *polo total* primer, was 2.18. In addition, the *CG6024 pA2* primer when used in the optimized condition of *polo pA2* primer, displayed an efficiency of 1.92.

After each primer pair's efficiency have been determined, qPCRs were performed to assess the levels of *CG6024* total and pA2 in *RpII215*, *w* ^{1118}, and *RpII140* flies. In Figure 10, *RpII215* mutant in comparison with the wild-type counterpart displayed comparable total mRNA levels of *CG6024* and a decrease in *CG6024* pA2, which are consistent with the results already published (Pinto et al. 2011). Interestingly, the *RpII140* flies in comparison with *w* ^{1118} showed a significantly decrease in total *CG6024* mRNA levels by 0.4 fold (i.e. 40%), and a significantly augmentation of *CG6024* pA2 levels by 3.3 fold (i.e. 330%). These results suggest that the *CG6024* short isoform *(CG6024* pA1*)* is reduced in *RpII140* flies, since *CG6024* total mRNA levels

decreased but *CG6024* pA2 levels increased. Additionally, these results also reveal a *CG6024* 3'UTR lengthening in these mutant flies.

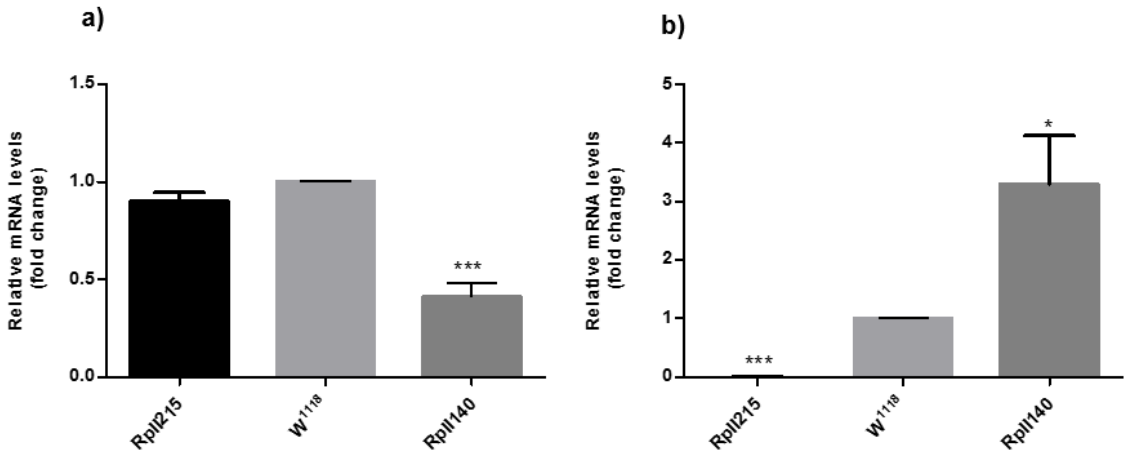

Figure 10- RNA Pol II kinetics with a slower and a faster transcription elongation rate affects *CG6024* alternative polyadenylation. a) *CG6024* mRNA levels in *RpII215, w^{1118}* and *RpII140* flies **b)** *CG6024* pA2 mRNA levels in *RpII215*, *w^{1118}* and *RpII140* flies. Graph represents fold change of p*olo* mRNAs in adult flies, relative to *7SL* mRNA. For all the panels, error bars show s.e.m. from at least three independent experiments

2. Study on the role of transcription elongation factors in *polo* alternative polyadenylation

To understand the impact of the elongation factors dELL and dSpt6 in alternative polyadenylation, each of these factors was depleted *via* RNAi from *Drosophila* S2 cells and then changes in *polo* poly(A) site usage were quantified.

2.1. Optimization of PCR products

The *in vitro* transcription procedure requires a purified DNA template produced by PCR. In this thesis, the DNA templates used were specific PCR products for each target gene: *dell*, *dspt6* and D*sRED*. These PCR products were synthesised using primers specific for the target genes flanked with T7 promoter regions, able to amplify the sequence of interest from S2 cells (*dell* and *dspt6* genes) or from the plasmid pIRES2 DsRed-EGFP (D*sRED* gene). In order to optimize the quality of the PCR products for each target gene, diverse PCR conditions were tested (primer and sample concentrations, PCR programs) and modifications in the purification method were performed.

Taking into account that some PCR products displayed subtle primer dimer bands, it was tested the usage of guanidine hydrochloride in the purification procedure, as suggested by the manufacturer. Figure 11 shows the result of the purification of *dell* PCR product with and

without guanidine hydrochloride, and both conditions display the expected band at 500bp and no primer dimer band.

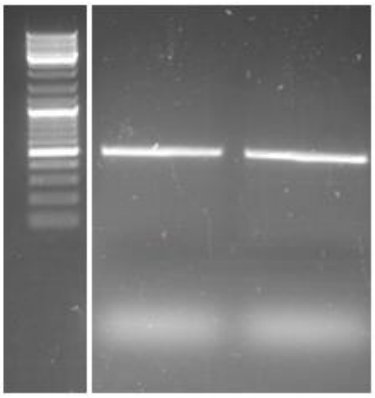

Figure 11- PCR product purification.
First lane is the molecular ladder of 100 to 1000bp. Purification using guanidine hydrochloride (lane on the left) and not (lane on the right).

Ultimately, all the mentioned optimization steps lead to PCR purified products without unspecific bands and primer dimers as illustrated in Figure 12. In fact, in this figure it is possible to observe the expected band of ~1000bp in the non-purified and purified PCR products of *dSpt6* and the expected band of ~500bp of non-purified and purified PCR products of d*ell* and *DsRED*. These PCR products were further used in *in vitro* transcription assays that resulted in the dsRNA shown in Figure 13 . The dsRNA for each target gene were displayed using 1µl and 3 µl of sample to assure that the band was visible. For all the dsRNAs it is possible to observe the expected band (~1000bp for *dspt6* and ~500bp for *dell* and *DsRED*).

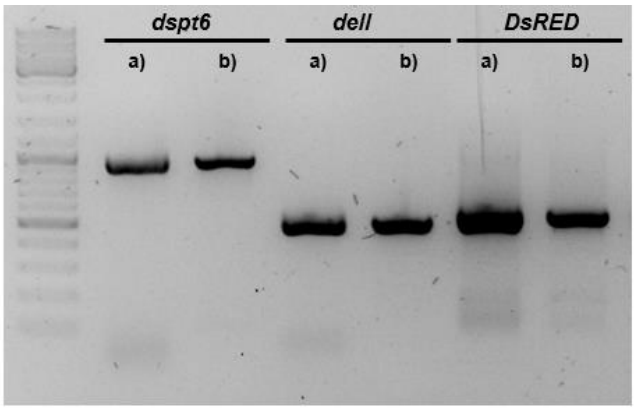

Figure 12 – Quality of PCR products.
First lane is the molecular ladder of 100 to 1000bp. **a)** before purification. **b)** after purification

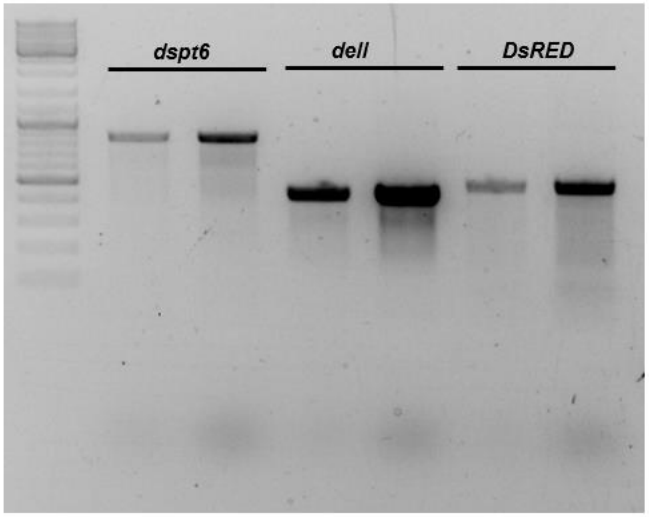

Figure 13-dsRNA after the *in vitro* transcription reaction.
First lane is the molecular ladder of 100 to 1000bp. dsRNA after reaction using 1µL (left lane) or 3 µL (right lane) of sample

2.2. Optimization of the RNAi assays

The knockdown assays here described were designed to be executed on exponentially growing S2 cells, therefore the proliferation state of the cells was analysed (Figure 14).

In Figure 14 it is possible to observe an initial phase from day 0 to day 2, in which the growing rate is moderated. Then, between the 2nd and 3rd day, an exponential phase begins with an enhanced proliferation rate that slows down in the last phase between the 3rd and 4th day.

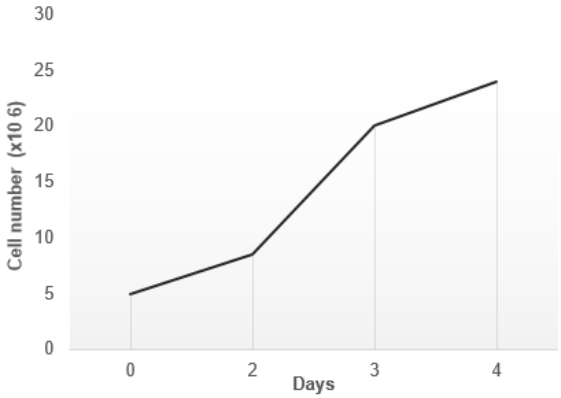

Figure 14- Graphic representation of S2 cells proliferation

After determining the S2 cells exponential growth phase, an optimization of the RNAi assays was in order. It was used 25 µg of each dsRNA and optimized the incubation period for each target.

To validate the knockdown efficiency in every experiment qPCR were performed, using primers to measure *dell* and *dspt6* mRNA levels in the dsRNA treated sample and control sample (treated with *DsRED* dsRNA). Both *dell* and *dspt6* primer pairs were optimized using 58°C as

annealing temperature and 0.15 µM of each primer, and the efficiency was 2.07 to *dell* primer pair and 1.9 *dspt6* primer pair.

The Figure 15 shows *dell* knockdown efficiency using different dsRNA incubation periods. The knockdown efficiency at 96h was 34% and at 120h was 40%, which means that the *dell* knockdown efficiency is greater at 120h. This is in agreement with the literature, since Smith *et al.* reached 50% of *dell* knockdown when using 2^6 cells/ml that were serum starved, treated with 3 boosts of 10 µg dsRNA and harvested at 120 h. Note that the dsRNA used in this work and the ones used by Smith *et al.* target different *dell* regions (Smith, Lin, Garrett, et al. 2011).

In order to improve the 40% efficiency, the dsRNA amount was rise up to 30 µg, which was added to the cells in 10 µg boosts at 0h, 48h and 96h (Figure 15). The knockdown efficiency did not improve, in fact no knockdown was achieved.

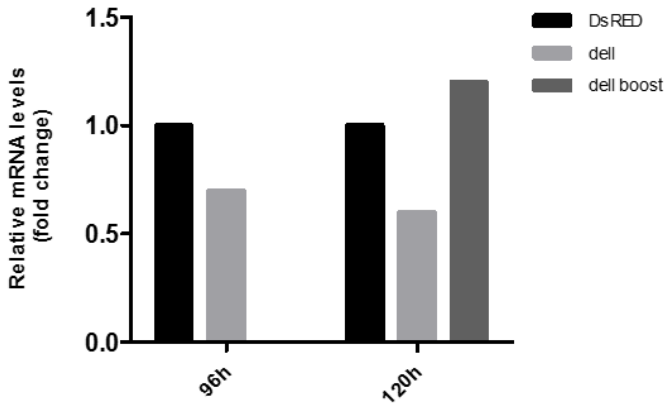

Figure 15- *dell* knockdown efficiency. Using different dsRNA incubation times (96h and 120h) and in single or boosted administrations. Graphs represents fold change of *dell* mRNAs in samples treated with *dell* or *DsRED* dsRNA, relative to *7SL* mRNA.

The *dspt6* knockdown efficiency at 96 and 120h is illustrated in Figure 16. The knockdown efficiency at 96h is 60% and at 120h is 2%. It is noteworthy that an increase of cell death was observed in the cells treated with *dspt6* dsRNA in comparison with the cells treated with *DsRED* (*data not shown*). In addition, the level of cell death was higher in the sample treated with *dspt6* dsRNA at 120h then at 96h.

The 60% *dspt6* knockdown efficiency is in agreement with the literature, since Ardehali *et al.* reached 70% of *dspt6* knockdown when using 1^6 cells/mL with serum free medium, treated with 10µg of dsRNA and harvested at 84h (Ardehali et al. 2009). Noteworthy, the dsRNA used in both works are distinct and target different *dspt6* regions (Ardehali et al. 2009).

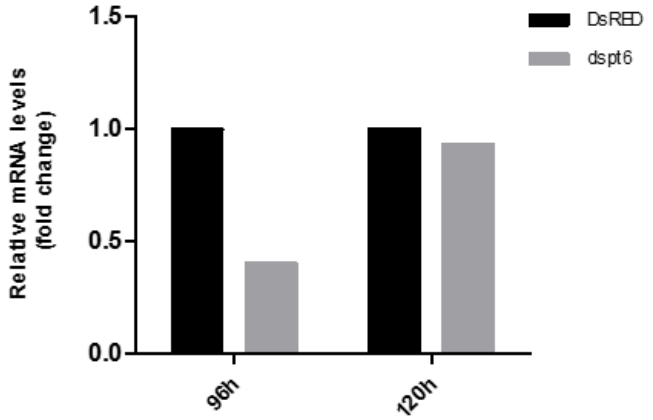

Figure 16- *dspt6* knockdown efficiency. Using dsRNA incubation times (96h and 120h). Graphs represents qPCR quantification of *dspt6* mRNAs in samples treated with *dspt6* or *DsRED* dsRNA, relative to 7SL mRNA.

2.3. Impact of elongation factors in polo alternative polyadenylation pattern

Once the knockdown of elongation factors where efficiently performed the effect in *polo* poly(a) site choice was assessed by qPCR.

a) dELL

Figure 17 displays total *polo* and pA2 mRNA levels in cells treated with *dell* and *DsRED* dsRNA for 120h. Similar *polo* total mRNA levels were observed in both samples (Figure 17a) indicating that this elongation factor does not affect polo mRNA levels. Interestingly, depletion of *dell* causes a 1.7 fold (i.e. 170% fold) increase in *polo* pA2 mRNA levels (Figure 17b)).

These results indicate that dELL modulates *polo* alternative polyadenylation and suggest that *polo* short isoform (*polo* pA1) is reduced when *dell* is diminished because *polo* total mRNA levels are the same but *polo* pA2 levels are increased. Therefore, dELL is enhancing the usage of *polo* proximal poly(A) site.

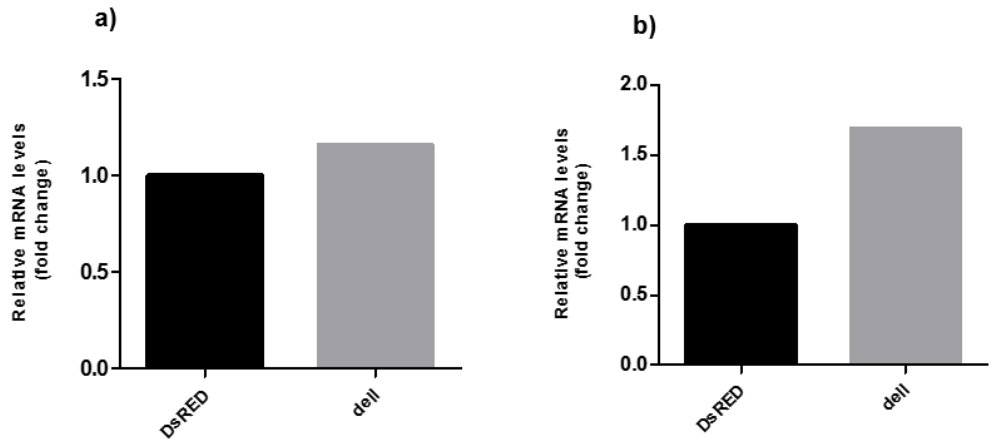

Figure 17- The elongation factor *dELL* affects *polo* alternative polyadenylation. a) *polo* total mRNA levels in samples treated with *dell* or *DsRED* dsRNA for 120h. **b)** *polo* pA2 mRNA levels in samples treated with *dell* or *DsRED* dsRNA for 120h. Graphs represent fold change of *polo* mRNAs in samples treated with each dsRNA relative to 7SL mRNA.

b) dSpt6

In Figure 18 it is displayed *polo* total and pA2 mRNA levels in samples treated with *dSpt6* and *DsRED* dsRNA for 96h. Similar *polo* total mRNA levels were detected in both samples (Figure 18a). Additionally, *polo* pA2 mRNA levels were different between samples, since the sample where *dspt6* was depleted displayed higher levels of *polo* pA2 than the control (6.8 fold) (Figure 18b).

Unfortunately, these results are not acceptable for a correct interpretation of the phenomena in study, since the expression of the reference gene used (*7SL*) was not invariant under the experimental conditions described (Bustin et al. 2009).

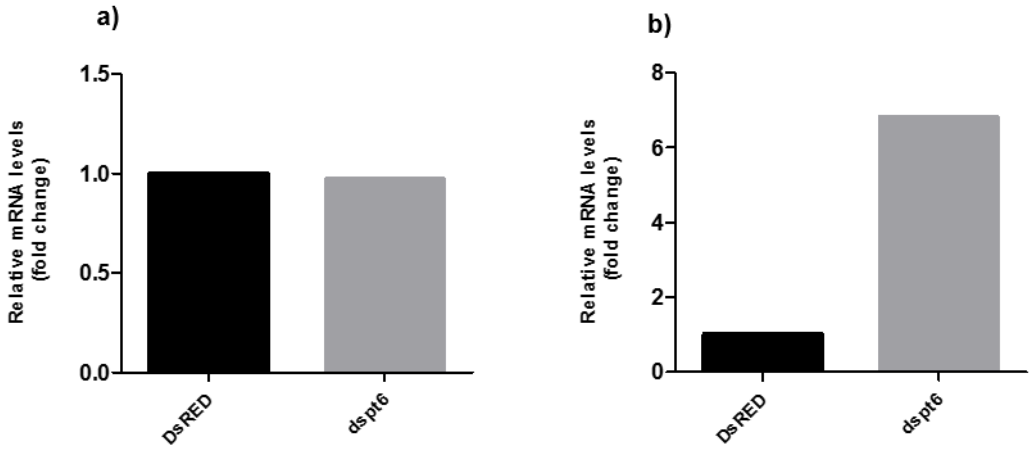

Figure 18- Elongation factor *dSpt6* affects *polo* alternative polyadenylation. a) *polo* total mRNA levels in samples treated with *dspt6* or *DsRED* dsRNA for 120h. **b)** *polo* pA2 mRNA levels in samples treated with *dspt6* or *DsRED* dsRNA for 120h. Graphs represent qPCR quantification of *polo* mRNAs in samples treated with each dsRNA, relative to *7SL* mRNA.

Chapter 5

Discussion and Future Perspectives

Alternative polyadenylation is a co-transcriptional event involved in the quantitative and qualitative regulation of gene expression in a diversity of biological processes. Through the selection of single or alternative poly(A) site in one transcription unit this process presents a major player influencing the dynamics of gene regulation (Millevoi & Vagner 2010; Shepard et al. 2011). In addition, transcription elongation is also described as capable of adjusting the outcome of gene expression, since it provides a mechanism for the modulation of cellular RNA levels (Kwak & Lis 2013; Svejstrup 2013). Interestingly, an intense and relevant cross-talk between both processes have been described (Luna et al. 2005).

The work presented in this thesis aims to understand the impact of an increased Pol II elongation rate in alternative polyadenylation. In addition, it was investigated the role of specific transcription elongation factors as relevant players.

In order to evaluate alternative polyadenylation in both models used in this project, the RNA was isolated using TRIzol, since this approach maintains the integrity of the RNA and is also particularly advantageous in situations where cells or tissues are enriched for endogenous RNases (Rio et al. 2010).

The RNA extraction condition in which RNA is not degraded and does not displayed genomic contamination was optimized. The presence of EDTA was important in the method, since samples treated with EDTA displayed lower levels of RNA degradation. The crucial role of EDTA is due the fact that this compound chelates divalent metals (primarily magnesium and calcium) which are co-factors and essential for DNase and RNAse activity. Additionally, it was shown that long DNase incubation periods display high levels of degradation, presumably because high temperatures can cause RNA denaturation and activate a few RNAses.

One of the aims of this project was the assessment of the alternative polyadenylation pattern in an enhanced Pol II rate scenario. In order to fulfil this goal, the flies used were mutants engineered to elongate at different speeds: 50% slower (*RpII215* flies) and 50% faster (*RpII140*

flies). Additionally, qPCRs were performed with DNA synthesised from RNA extracted from 20 male flies. It should be noted that only males were used in order to discard the variability that would arise from the usage of female flies, since sex-specific expression of alternative transcripts in *Drosophila* is a well-known phenomenon, in which gene expression profiles are dominated by the reproductive differences between males and females (McIntyre et al. 2006; Smibert et al. 2012).

The *RpII140* flies in comparison with the wild-type showed a decrease in *polo* levels and similar *polo* pA2 mRNA levels, which suggests that *polo* pA1 is decreased. Interestingly, this effect is observed on the stronger poly(A) site (pA1).

To better understand these results and assess the role of the strength of the poly(A) signals in the mechanism, the *CG6024* gene was used. This gene displays a similar genomic structure with *polo*, as it contains two poly(a) sites in the 3'UTR, however the strength of the poly(A) signals are opposite to the one exhibit by *polo* . Indeed, in *CG6024* pA2 display a canonical signal (AATAAA) and pA1 shows a variant of it, consequently pA2 signal is stronger than pA1 (information kindly provided by Joel Graber and GenBank, NCBI) .

The *RpII140* flies in comparison with w^{1118} showed a decrease in *CG6024* total mRNA levels and an increase of *CG6024* pA2 levels, which suggest that *CG6024* pA1 is reduced. These results also reveal a *CG6024* 3'UTR lengthening in these mutant flies.

The results here presented may constitute preliminary evidence that the decrease of total mRNA levels is a general event when the Pol II processivity is enhanced, which presumably can be due a decrease in the mRNA 3'end processing efficiency (either or in both cleavage and polyadenylation) (Proposed working model in Figure 19). We propose that a decrease in the mRNA 3'end formation efficiency can be a direct consequence of the enhancement of the elongation rate. In that scenario, the efficiency of nucleotide selection and incorporation and thus the nascent RNA production is faster. It is possible then that the recognition of the poly(A) signal - which is necessary for transcription termination - is affected, which could lead to an increased Pol II read-through and consequently a decrease in total mRNA levels. In fact, a similar relationship between elongation rate and splicing efficiency have been described by Braberg *et al.*, since it was proved in yeast that fast Pol II mutations result in diminished splicing efficiency (Braberg et al. 2013; Moehle et al. 2014). The work here presented suggests that some valid and similar conclusions can also be applied to mRNA 3'end processing (cleavage and alternative polyadenylation).

In *RpII140* flies, the decrease of total mRNA levels may also be due to an indirect effect of RNA quality control mechanism. To clarify this question, future studies should be performed to assess the expression of RNA quality control genes in *RpII140* flies in comparison to wild-type. Using qPCR it could be assessed the expression of genes such as Upf1, which is crucial to RNA quality control because it is part of the so-called "decay-inducing complex" that targets mRNA for decay (Isken et al. 2008; Rehwinkel et al. 2005; Isken & Maquat 2007). If the expression of Upf1

is enhanced in *RpII140* flies that could suggest that the decrease of total mRNA levels is due a regulatory mechanism of RNA quality control.

For every transcript total RNA levels are the result of a fine-tuned balance between RNA synthesis and RNA decay, and the results here shown reflect what happens at a steady state. Therefore, it would be important in the future to analyse nascent transcripts, since analysis performed with total cellular mRNA show results with poor resolution of the underlying mechanisms of the alterations in cellular response and a bias towards detecting up-regulation of short-lived transcripts.

An enhanced Pol II elongation rate leads to a subtle decrease in the usage of the proximal site (pA1) in *polo* and a more significant one in *CG6024*. Additionally, in the *CG6024* gene the distal site is preferentially used in *RpII140* flies, which indicates that there is a shift in the alternative polyadenylation pattern. Presumably, the statistical significant shift in the poly(A) site usage in *CG6024* is due to the distal signal (pA2) higher strength, and the subtle pA1 reduction in *polo* is due the weakness of the distal signal.

The different outcomes in both genes studied also suggest that in addition to signal strength other regulatory elements (e.g. USE, DSE, various polyadenylation factors) can affect poly(A) site choice in an enhanced Pol II elongation rate scenario. The relevance of such elements in alternative polyadenylation was proved by Hall-Pogar *et al.*, by showing that the synthesis of both *cox-2*mRNAs seems to be tissue-specific and that the expression of the shorter mRNA is dependent of an USE near the proximal poly(A) site (Hall-Pogar et al. 2005). Moreover, *polo* contains an USE that affects APA (Moreira and coll., unpublished data). Therefore, future studies must be performed with other genes that in addition to the two poly(A) sites in the 3'UTR, should also exhibit similar regulatory elements (e.g USE and DSE). This would allow to assess if these elements are key in the modification of the alternative polyadenylation pattern when the Pol II processivity is enhanced. The selection of such genes could be performed bioinformatically, by performing a RNA-sequencing (RNA-seq). Using the gene selected and mutating their USE region, qPCRs studies could be performed to assess if the presence of e.g. USE correlates with alternative polyadenylation alterations when the Pol II rate is changed.

It is possible that the point mutation in Pol II in the *RpII215* and *RpII140* flies alter interactions with specific factors that are differently recruited to the elongating Pol II and thus affect the alternative polyadenylation event. A pertinent study would be to assess differences in the recruitment of some key factors in polyadenylation, in both mutants. It could be interesting to perform chromatin immunoprecipitation (Co-Ip) assays, with antibodies against Pol II and e.g. one CPSF's subunits, since this factor is required for poly(A) signal recognition, endonucleotic cleavage and polyadenylation reactions. If alternative polyadenylation effects observed in *RpII140* were strictly due differences in the elongation rate, one would expect this mutant to have similar levels of CPSF recruited to the Pol II than wild-type and *RpII215*.

To the best of our knowledge, this study represents the first example where an enhanced Pol II transcription elongation rate displays an impact in the mRNA 3'end formation efficiency and alternative polyadenylation pattern. It is noteworthy that in splicing it is described that not all genes respond to Pol II transcription rate modifications (Moehle et al. 2014), which can also be the case for alternative polyadenylation. Even though it is technically difficult to measure gene-specific transcription elongation rates *in vivo*, it would probably be found that genes affected by Pol II rate already would be genes in which elongation rate is modified. Alternatively, perhaps genes that did not respond to changes in Pol II rate may have a powerful mechanism for maintaining normal Pol II elongation rate even in the context of the Pol II mutations, to assure the maintenance of protein levels (Moehle et al. 2014).

Taken together, all these evidences suggest that kinetic coupling of transcription elongation and alternative polyadenylation is indeed a relevant regulatory pathway of gene expression.

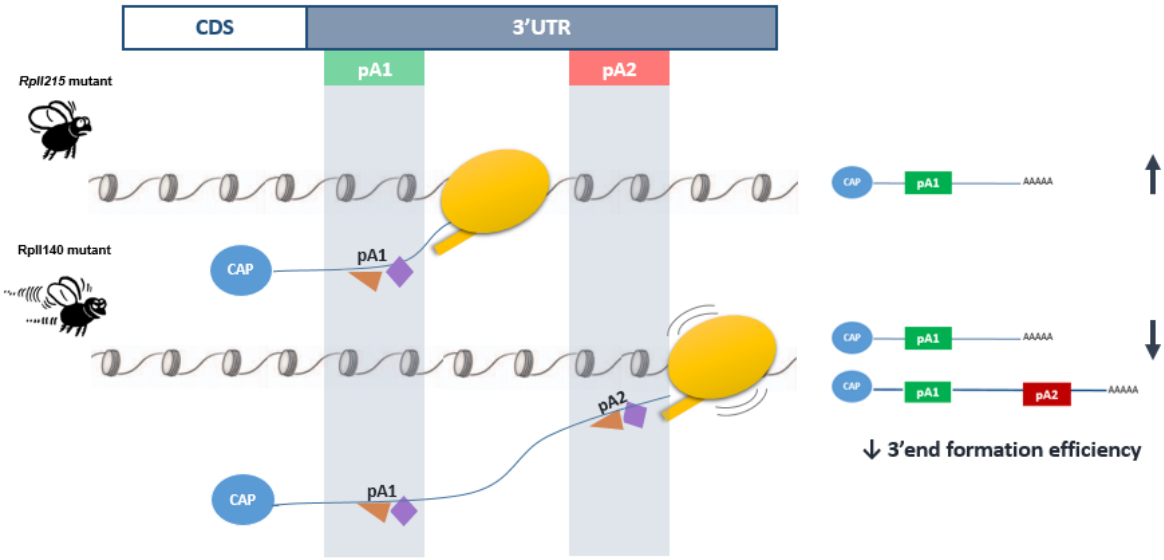

Figure 19- Hypothetical model of how Pol II elongation rate impacts 3'end formation efficiency. (See text for details)

The results presented in this thesis elucidate the function of elongation factors in alternative polyadenylation. Based on the present literature, dELL and dSpt6 were the elongation factors selected to study and RNAi methodology was performed to knockdown each of these factors from *Drosophila* S2 cells, in order to assess their effect in *polo* alternative polyadenylation pattern.

One of the elongation factors studied in this project was dELL, which increases the transcription elongation rate of Pol II, participates in SEC and LEC and also has a widespread role in gene transcription during development (Smith, Lin & Shilatifard 2011; Smith, Lin, Garrett,

et al. 2011). It should be highlighted that *dell* was previously shown to be an essential gene in *Drosophila*, and homozygous for loss-of-function alleles die at the end of embryogenesis or in early first instar larvae, therefore the RNAi technique here performed was important to assess dELL role in alternative polyadenylation (Eissenberg et al. 2002). In fact, knockdown *dell* by RNAi, typically reduces, but does not eliminate, the targeted gene products. It would be interesting to perform tissue-specific RNAi and/or overexpression of *dell* in flies to investigate its effect in alternative polyadenylation in the context of the whole organism.

In order to accomplish the strongest *dell* depletion possible, various knockdown conditions were tested. It was used 30µg and boosted administrations in order to enhance the knockdown efficiency, however the new conditions did not lead to *dell* depletion, probably because the dsRNA was degraded when the assays were performed, since RNA samples are more prone to degradation that DNA samples. In fact, by using this approach it would be expected that the knockdown efficiency achieved would be equal or higher than 40%, since higher dsRNA amount was used and it was administrated at 3 distinct time points.

Our results suggest that dELL enhances the usage of *polo* proximal poly(A) site. Similarly, these results are in agreement with the work performed in dELL mammalian homolog ELL2, which suggests a conservation of the mechanisms. It was shown that ELL2 enhance usage of the proximal poly(A) site of IgM and it was proposed that the loading of ELL2 and CstF-64 on Pol II was interligated and it caused that effect in proximal poly(A) signal usage (Martincic et al. 2009).

The impact of dELL in poly(A) site usage could presumably be due to a direct action of this factor in alternative polyadenylation , which suggests a novel role for dELL as a relevant player in this mechanism. Nevertheless, dELL exhibit various roles that can influence the results obtained, therefore it is plausible that the correspondence between *dell* knockdown and the impact in alternative polyadenylation is due to an indirect effect and not a direct action of this factor on the pre-mRNA. Indeed, dELL participates in SEC, which probably can explain the described outcome: when *dell* is reduced SEC assembly is compromised and Pol II is destabilized, which may affect the poly(A) site choice (Byun et al. 2012). Other possibility, is concerning dELL's role in the Pol II elongation rate, therefore when *dell* is knockdown the Pol II elongation rate is affected and that influences poly(A) site usage, as previously shown. Finally, dELL is also suggested to promote loading of various polyadenylation factors as CstF-64 on the transcription machinery, which can enhance the usage of the proximal poly(A) site of *polo*. This is accordant to what was published by Martincic *et al.* regarding the ELL2, since it was suggested that this factor influences the poly(A) site choice of IgM by affecting the loading of CstF-64 onto Pol II (Martincic et al. 2009).

In order to clarify the direct / indirect role of dELL in poly(A) site choice, further studies should be performed. To assess if dELL presence in SEC is affecting alternative polyadenylation, depletion of other SEC components (e.g. Lilli) using RNAi could be performed. If the same

tendency was obtained then probably dELL acts in poly(A) site choice indirectly as a SEC member (Luo et al. 2012). On the other hand, if dELL impacts on the elongation rate *per se* and this is responsible for alternative polyadenylation alterations, then the reduction of another elongation factor that stimulates positively Pol II rate, such as TFIIF, would lead to a similar outcome as the one here observed (Bengal et al. 1991; Tan et al. 1995). Additionally, if the rational of ELL2 described by *Martincic et al.* also applies for dELL, then this factor acts indirectly in poly(A) site choice by affecting the loading of CstF-64 (Martincic et al. 2009). To prove that, CstF-64 could be depleted by RNAi assays and if similar results to *dell* knockdown were obtained, then dELL influences indirectly alternative polyadenylation by affecting the loading of CstF-64. Furthermore, to verify that the effect described in this document is a general effect more genes should be studied and ultimately high-throughput RNA-seq methodologies should be employed.

The findings reported in this thesis show that dELL is one of the few transcription elongation factors described as enhancing promoter proximal poly(A) site 'choice'. Therefore, the results here presented were incorporated into a working model (Figure 20).

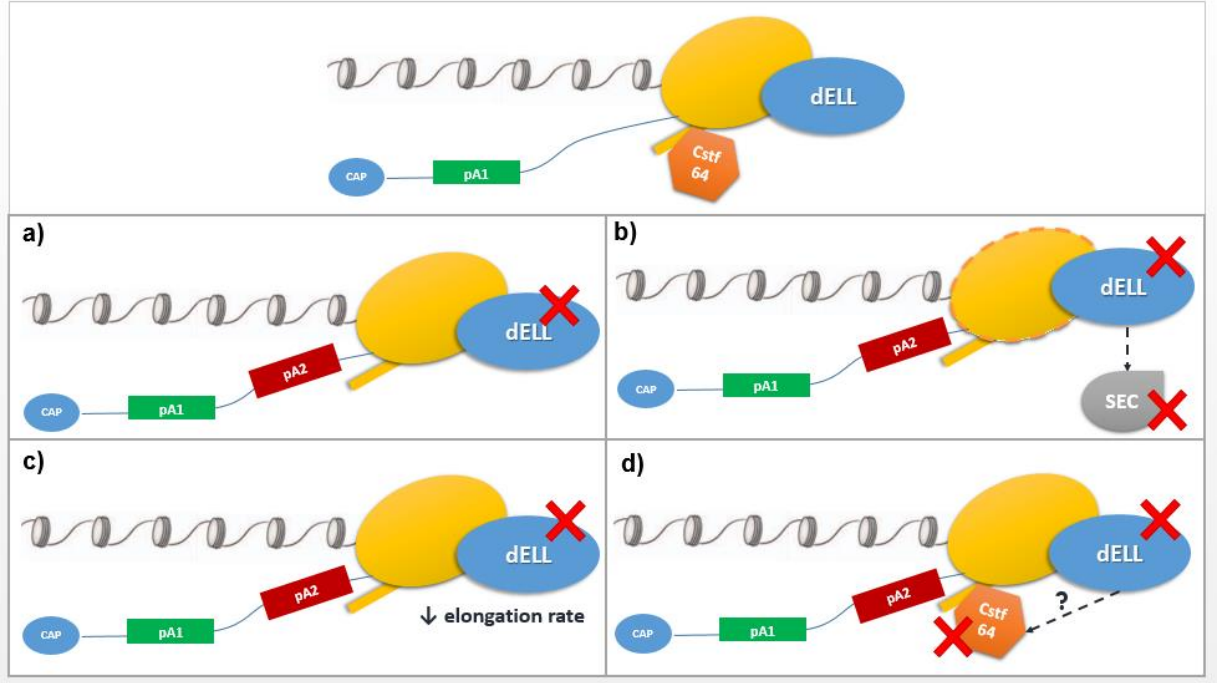

Figure 20- Working models on the role of dELL in poly(A) site choice. Upper panel: in dELL presence scenario **Lower panel:** in dELL depletion scenario **a)** absence of dELL have a direct impact in alternative polyadenylation **b)** lack of dELL affects SEC assembly and consequently Pol II stabilization **c)** dELL depletion can reduce Pol II elongation rate **d)** dELL reduction disturbs Cstf64 loading to CTD**.** (See text for details)

The other elongation factor studied in this project was dSpt6, which positively stimulates the elongation rate of Pol II, regulates histone modifications and is critical to normal transcription termination (Ardehali et al. 2009).

As previously mentioned, the *dspt6* knockdown affected the cells viability which highlight the relevance of this elongation factor for cell survival. Unfortunately, from the data presented nothing can be extrapolated since the expression of the reference gene (7SL) fluctuated between the samples, which is not the recommended behaviour (Bustin et al. 2009). This variation may be due to cell death already mentioned and in the future, these studies should be repeated using other reference gene, e.g. *rp49* gene.

Regardless the outcome of these future studies, those results must be carefully analysed because a possible effect of *dspt6* knockdown in alternative polyadenylation, may be due a direct action of this factor but also to its role in termination or as a histone modifier. Indeed, *dspt6* depletion will probably affect the chromatin structure, since dSpt6 remove nucleosome barriers by direct interaction with histones H3 and H4. Therefore, dSpt6 absence could affect Pol II action by making the DNA template unaccessible for productive elongation, which may lead to modifications in alternative polyadenylation (Kaplan et al. 2005). To assess if impact of *dspt6* depletion in poly(A) site usage is due to its role as histone modifier, another histone modifiers should be tested.

In addition, *dspt6* silencing can also affect the termination process. Therefore, it should be investigated the Pol II distribution along the *polo* gene when *dspt6* is depleted by performing Chromatin Immunoprecipitation (ChIP) using e.g. α-Rpb3 Pol II antibody. Transcription termination in high eukaryotes largely depends on the presence of a poly(A) signal. When Pol II terminates transcription it disengages from the DNA template and this effect is observed by a decrease in Pol II occupancy in the 3'UTR, downstream of the poly(A) signal. Therefore, it will be possible to asses if the alternative polyadenylation outcome can be due to the termination defect and not due a direct action of dSpt6. Additionally, *spt6* mutations are described to impair transcription read-through (Kaplan et al. 2005). Since *snap* is *polo* downstream gene, *snap* levels must be also assessed by nuclear run-on, which measures nascent *spt6* mRNAs, *polo* mRNAs and *snap* mRNA readthrough transcripts. If different levels are obtained when *dspt6* is depleted, then probably the impact of dSpt6 in poly(A) site choice is due to a defect in transcription read-through.

The RNAi studies performed to assess the role of dELL and dSpt6 in alternative polyadenylation, should be also complemented with western assays to confirm the decrease in the protein levels (Professor Ali Shilatifard and Professor John Lis kindly provided dELL and dSpt6 antibodies, respectively). In addition, it would be interesting to perform *in vivo* studies, using RNAi mutants for each named factor and assess if there is an interesting outcome that would complement the *in vitro* studies.

As previously mentioned, the coexistence of both kinetic and recruitment approaches has been proposed in an integrated model. In the context of this project, an interesting experiment to perform would be to assess if there is a differential recruitment of elongation factors (e.g. dELL and dSpt6) in the *RpII215* and *RpII140* flies mutants, by Co-IP and ChIP assays. Hopefully, this and other experiments will give more insight in the bidirectional crosstalk between the kinetic and recruitment model in alternative polyadenylation, by elucidating the relevance of the interaction of elongation rates and elongation factors in alternative polyadenylation.

Many emerging concepts justify the renewed interest in the cross-talk between transcriptional elongation and alternative polyadenylation. We hope that the results here presented highlighted that alternative polyadenylation and transcription elongation can be recognized as key mechanisms for eukaryotic gene regulation.

In the future, the knowledge obtained from how alternative polyadenylation and elongation interact, will certainly help to gain a novel insight on the molecular mechanisms involved in regulation of gene expression, and hopefully also to develop new tools to understand, detect and treat some human disease.

References

Adelman, K. et al., 2009. Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling. *Proceedings of the National Academy of Sciences of the United States of America*, 106(43), pp.18207–12.

Aebi, M. & Weissman, C., 1987. Precision and orderliness in splicing. *Trends in Genetics*, 3, pp.102–107.

Alberts, B. et al., 2008. *Molecular Biology of the Cell* 5th ed., Garland Press.

Andreassi, C. & Riccio, A., 2009. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends in cell biology*, 19(9), pp.465–74.

Andrulis, E.D. et al., 2002. The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. *Nature*, 420(6917), pp.837–41.

Ardehali, M.B. et al., 2009. Spt6 enhances the elongation rate of RNA polymerase II in vivo. *The EMBO journal*, 28(8), pp.1067–77.

Ashe, M.P. et al., 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes & Development*, 9(23), pp.3008–3025.

Aso, T. et al., 1995. Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. *Science (New York, N.Y.)*, 269(5229), pp.1439–43.

Banerjee, A. et al., 2013. PABPN1: molecular function and muscle disease. *The FEBS journal*, 280(17), pp.4230–50.

Bengal, E. et al., 1991. Role of the Mammalian Transcription Factors IIF , IIS , and IIX during Elongation by RNA Polymerase II. *Molecular and cellular biology*, 11(3).

Bentley, D.L., 2014. Coupling mRNA processing with transcription in time and space. *Nature reviews. Genetics*, 15(3), pp.163–75.

Berg, M.G. et al., 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell*, 150(1), pp.53–64.

Bio-Rad Laboratories, 2006. Real-Time PCR Applications Guide.

Boelens, W.C. et al., 1993. The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell*, 72(6), pp.881–92.

Boettiger, A.N. & Levine, M., 2009. Synchronous and stochastic patterns of gene activation in the Drosophila embryo. *Science (New York, N.Y.)*, 325(5939), pp.471–3.

Bortvin, A. & Winston, F., 1996. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science*, 272(5267), pp.1473–6.

Braberg, H. et al., 2013. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell*, 154(4), pp.775–88.

Brais, B. et al., 1998. Short GCG expansions in PABP2 gene cause oculopharyngeal muscular dystrpphy. *Nature Genetics*, 18, pp.164–167.

Buratowski, S., 2005. Connections between mRNA 3' end processing and transcription termination. *Current opinion in cell biology*, 17(3), pp.257–61.

Bustin, S. a et al., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*, 55(4), pp.611–22.

Byun, J.S. et al., 2012. ELL facilitates RNA polymerase II pause site entry and release. *Nature communications*, 3, p.633.

Castelo-branco, P. et al., 2004. Polypyrimidine Tract Binding Protein Modulates Efficiency of Polyadenylation. *Molecular and cellular biology*, 24(10).

Castro, M.A.A. et al., 2007. Extracellular isoforms of CD6 generated by alternative splicing regulate targeting of CD6 to the immunological synapse. *Journal of immunology*, 178(7), pp.4351–61.

Chen, Y. et al., 1996. Drosophila RNA polymerase II mutants that affect transcription elongation. *Journal of Biological Chemistry*, 271(11), pp.5993–9.

Chen, Y. et al., 1993. Mapping mutations in genes encoding the two large subunits of Drosophila RNA polymerase II defines domains essential for basic transcription functions and for proper expression of developmental genes. *Molecular and cellular biology*, 13(7), pp.4214–22.

Chopra, V.S., Hong, J.-W. & Levine, M., 2009. Regulation of Hox gene activity by transcriptional elongation in Drosophila. *Current biology*, 19(8), pp.688–93.

Close, P. et al., 2012. DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. *Nature*, 484(7394), pp.386–9.

Colgan, D.F. & Manley, J.L., 1997. Mechanism and regulation of mRNA polyadenylation. *Genes & Development*, 11(21), pp.2755–2766.

Cougot, N. et al., 2004. "Cap-tabolism". *Trends in biochemical sciences*, 29(8), pp.436–44.

Coulter, D.E. & Greenleaf, a L., 1985. A mutation in the largest subunit of RNA polymerase II alters RNA chain elongation in vitro. *The Journal of biological chemistry*, 260(24), pp.13190–8.

Cui, Y. & Denis, C.L., 2003. In Vivo Evidence that Defects in the Transcriptional Elongation Factors RPB2, TFIIS, and SPT5 Enhance Upstream Poly(A) Site Utilization. *Molecular and Cellular Biology*, 23(21), pp.7887–7901.

Curinha, A. et al., 2014. Implications of polyadenylation in health and disease. *Nucleus*, 5(6), pp.1–12.

Danckwardt, S. et al., 2011. p38 MAPK controls prothrombin expression by regulated RNA 3' end processing. *Molecular cell*, 41(3), pp.298–310.

Danckwardt, S., Hentze, M.W. & Kulozik, A.E., 2008. 3' end mRNA processing: molecular mechanisms and implications for health and disease. *The EMBO journal*, 27(3), pp.482–98.

Darnell, J.E., 2013. Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. *RNA (New York, N.Y.)*, 19(4), pp.443–60.

Derti, A. et al., 2012. A quantitative atlas of polyadenylation in five mammals. *Genome research*, 22, pp.1173–1183.

Dillman, A. a et al., 2013. mRNA expression, splicing and editing in the embryonic and adult mouse cerebral cortex. *Nature neuroscience*, 16(4), pp.499–506.

Dronamraju, R. & Strahl, B.D., 2014. A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic acids research*, 42(2), pp.870–81.

Dujardin, G. et al., 2014. How slow RNA polymerase II elongation favors alternative exon skipping. *Molecular cell*, 54(4), pp.683–90.

Dujardin, G. et al., 2013. Transcriptional elongation and alternative splicing. *Biochimica et biophysica acta*, 1829(1), pp.134–40.

Dürr, J. et al., 2014. The transcript elongation factor SPT4/SPT5 is involved in auxin-related gene expression in Arabidopsis. *Nucleic acids research*, 42(7), pp.4332–47.

Egloff, S., Dienstbier, M. & Murphy, S., 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends in genetics : TIG*, 28(7), pp.333–41.

Eissenberg, J.C. et al., 2002. dELL is an essential RNA polymerase II elongation factor with a general role in development. *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp.9894–9.

Elkon, R., Ugalde, A.P. & Agami, R., 2013. Alternative cleavage and polyadenylation: extent, regulation and function. *Nature reviews. Genetics*, 14(7), pp.496–506.

Fabian, M.R., Sonenberg, N. & Filipowicz, W., 2010. Regulation of mRNA translation and stability by microRNAs. *Annual review of biochemistry*, 79, pp.351–79.

Fleige, S. & Pfaffl, M.W., 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular aspects of medicine*, 27(2-3), pp.126–39.

Garncarz, W. et al., 2013. A high-throughput screen to identify enhancers of ADAR-mediated RNA-editing. *RNA biology*, 10(2), pp.192–204.

Gehring, N.H. et al., 2001. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. *Nature genetics*, 28(4), pp.389–92.

Gerber, M. et al., 2001. Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo. *The EMBO journal*, 20(21), pp.6104–14.

Di Giammartino, D.C., Nishida, K. & Manley, J.L., 2011. Mechanisms and consequences of alternative polyadenylation. *Molecular cell*, 43(6), pp.853–66.

Gilchrist, D. a et al., 2010. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell*, 143(4), pp.540–51.

Da Glória, V.G. et al., 2014. T cell activation regulates CD6 alternative splicing by transcription dynamics and SRSF1. *Journal of immunology*, 193(1), pp.391–9.

Gruber, A.R. et al., 2013. Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley interdisciplinary reviews. RNA*.

Guhaniyogi, J. & Brewer, G., 2001. Regulation of mRNA stability in mammalian cells. *Gene*, 265(1-2), pp.11–23.

Gunderson, S.I. et al., 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell*, 76(3), pp.531–541.

Hall-Pogar, T. et al., 2005. Alternative polyadenylation of cyclooxygenase-2. *Nucleic acids research*, 33(8), pp.2565–79.

He, N. et al., 2011. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *PNAS*, 108(36).

Heidemann, M. et al., 2013. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochimica et biophysica acta*, 1829(1), pp.55–62.

Henriques, T., 2011. *Integrating transcriptional termination with gene expression of polo*. Universidade do Porto.

Hintermair, C. et al., 2012. Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *The EMBO journal*, 31(12), pp.2784–97.

Howe, K.J., 2002. RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochimica et biophysica acta*, 1577(2), pp.308–24.

Hsin, J.-P. & Manley, J.L., 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & development*, 26(19), pp.2119–37.

Hu, J. et al., 2005. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA*, 11(10), pp.1485–93.

Isken, O. et al., 2008. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell*, 133(2), pp.314–27.

Isken, O. & Maquat, L.E., 2007. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes & development*, 21(15), pp.1833–56.

Jalkanen, A.L., Coleman, S.J. & Wilusz, J., 2014. Determinants and implications of mRNA poly(A) tail size - Does this protein make my tail look big? *Seminars in cell & developmental biology*.

Jennings, B.H., 2013. Pausing for thought: disrupting the early transcription elongation checkpoint leads to developmental defects and tumourigenesis. *BioEssays*, 35(6), pp.553–60.

Ji, Z. et al., 2011. Transcriptional activity regulates alternative cleavage and polyadenylation. *Molecular systems biology*, 7(534), p.534.

Jin, Y. et al., 2007. RNA editing and alternative splicing of the insect nAChR subunit alpha6 transcript: evolutionary conservation, divergence and regulation. *BMC evolutionary biology*, 7, p.98.

Kaplan, C.D., Holland, M.J. & Winston, F., 2005. Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. *The Journal of biological chemistry*, 280(2), pp.913–22.

Kato, H., Okazaki, K. & Urano, T., 2013. Spt6: two fundamentally distinct functions in the regulation of histone modification. *Epigenetics : official journal of the DNA Methylation Society*, 8(12), pp.1249–53.

Kettenberger, H., Armache, K. & Cramer, P., 2004. Structure and Its Interactions with NTP and TFIIS. *Molecular cell*, 16, pp.955–965.

Kim, J.B. & Sharp, P. a, 2001. Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *The Journal of biological chemistry*, 276(15), pp.12317–23.

Kim, M. et al., 2004. Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *The EMBO journal*, 23(2), pp.354–64.

Kim, W.J., Burke, L.P. & Mortin, M.A., 1994. Molecular modeling of RNA polymerase II mutations onto DNA polymerase I. *Journal of molecular biology*, 244(1), pp.13–22.

Kornblihtt, A.R. et al., 2013. Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nature reviews. Molecular cell biology*, 14(3), pp.153–65.

Krogan, N.J. et al., 2003. Methylation of Histone H3 by Set2 in Saccharomyces cerevisiae Is Linked to Transcriptional Elongation by RNA Polymerase II. *Molecular and cellular biology*, 23(12).

Kwak, H. & Lis, J.T., 2013. Control of transcriptional elongation. *Annual review of genetics*, 47(September), pp.483–508.

De la Mata, M. et al., 2003. A slow RNA Polymerase II affects Alternative Splicing in vivo. *Molecular cell*, 12, pp.525–532.

Lenasi, T. & Barboric, M., 2010. P-TEFb stimulates transcription elongation and pre-mRNA splicing through multilateral mechanisms. *RNA Biology*, (April), pp.145–150.

Lewis, B.P., Green, R.E. & Brenner, S.E., 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1), pp.189–92.

Lewis, J.D. & Izaurralde, E., 1997. The role of the cap structure in RNA processing and nuclear export. *European journal of biochemistry / FEBS*, 247(2), pp.461–9.

Li, X.-Y. & Green, M.R., 1996. Tumorigenesis: Transcriptional elongation and cancer. *Current Biology*, 6(8), pp.943–944.

Licatalosi, D.D. et al., 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, 456(7221), pp.464–9.

Lidschreiber, M., Leike, K. & Cramer, P., 2013. Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Molecular and cellular biology*, 33(19), pp.3805–16.

Lin, C. et al., 2010. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Molecular cell*, 37(3), pp.429–37.

Lin, C. et al., 2013. The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation. *Cell*, 152(1-2), pp.144–56.

Liu, X., Bushnell, D. a & Kornberg, R.D., 2013. RNA polymerase II transcription: structure and mechanism. *Biochimica et biophysica acta*, 1829(1), pp.2–8.

Livak, K.J. & Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25(4), pp.402–8.

Llamazares, S. et al., 1991. polo encodes a protein kinase homolog required for mitosis in Drosophila. *Genes & Development*, 5(12a), pp.2153–2165.

Lou, H. et al., 1998. Regulation of alternative polyadenylation by U1 snRNPs and SRp20. *Molecular and cellular biology*, 18(9), pp.4977–85.

Luna, R. et al., 2005. Interdependence between transcription and mRNP processing and export, and its impact on genetic stability. *Molecular cell*, 18(6), pp.711–22.

Luo, Z., Lin, C. & Shilatifard, A., 2012. The super elongation complex (SEC) family in transcriptional control. *Nature Publishing Group*, 13(9), pp.543–547.

Lutz, C.S., 2008. Alternative polyadenylation: a twist on mRNA 3' end formation. *ACS chemical biology*, 3(10), pp.609–17.

Lutz, C.S. & Moreira, A., 2010. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. *Wiley interdisciplinary reviews. RNA*, 2(1), pp.22–31.

Mandal, S.S. et al., 2004. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*, 101(20), pp.7572–7.

Mapendano, C.K. et al., 2010. Crosstalk between mRNA 3' end processing and transcription initiation. *Molecular cell*, 40(3), pp.410–22.

Marshall, N.F. & Price, D.H., 1992. Control of Formation of Two Distinct Classes of RNA Polymerase II Elongation Complexes. *Mol. Cell. Biol.*, 12(5).

Martincic, K. et al., 2009. Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nature immunology*, 10(10), pp.1102–9.

Mayr, C. & Bartel, D.P., 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*, 138(4), pp.673–84.

McCracken, S. et al., 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, 385(6614), pp.357–61.

McIntyre, L.M. et al., 2006. Sex-specific expression of alternative transcripts in Drosophila. *Genome biology*, 7(8), p.R79.

Millevoi, S. & Vagner, S., 2010. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic acids research*, 38(9), pp.2757–74.

Mischo, H.E. & Proudfoot, N.J., 2013. Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochimica et biophysica acta*, 1829(1), pp.174–85.

Moehle, E.A. et al., 2014. Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA biology*, 11(4), pp.313–9.

Montecucco, A. & Biamonti, G., 2013. Pre-mRNA processing factors meet the DNA damage response. *Frontiers in genetics*, 4(June), p.102.

Montes, M. et al., 2012. TCERG1 regulates alternative splicing of the Bcl-x gene by modulating the rate of RNA polymerase II transcription. *Molecular and cellular biology*, 32(4), pp.751–62.

Montgomery, M.K., 2004. The use of double-stranded RNA to knock down specific gene activity. *Methods in molecular biology*, 260, pp.129–44.

Moore, M.J. & Proudfoot, N.J., 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*, 136(4), pp.688–700.

Moreira, A., 2011. Integrating transcription kinetics with alternative polyadenylation and cell cycle control. *Nucleus (Austin, Tex.)*, 2(6), pp.556–61.

Moreira, A. et al., 1995. Upstream sequence elements enhance poly(A) site efficiency of the C2 complement gene and are phylogenetically conserved. *The EMBO journal*, 14(15), pp.3809–19.

Nagaike, T. et al., 2011. Transcriptional activators enhance polyadenylation of mRNA precursors. *Molecular cell*, 41(4), pp.409–418.

Narita, T. et al., 2003. Human Transcription Elongation Factor NELF : Identification of Novel Subunits and Reconstitution of the Functionally Active Complex. *Molecular and cellular biology*, 23(6).

Nechaev, S. & Adelman, K., 2011. Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochimica et biophysica acta*, 1809(1), pp.34–45.

Neugebauer, K.M., 2002. On the importance of being co-transcriptional. *Journal of Cell Science*, 115(20), pp.3865–3871.

Nogués, G. et al., 2003. Control of alternative pre-mRNA splicing by RNA Pol II elongation: faster is not always better. *IUBMB life*, 55(4-5), pp.235–41.

Palangat, M. & Larson, D.R., 2012. Complexity of RNA polymerase II elongation dynamics. *Biochimica et biophysica acta*, 1819(7), pp.667–72.

Pfaffl, M., 2004. Quantification strategies in real-time PCR. In S. A. Bustin, ed. *A-Z of quantitative PCR*. International University Line, pp. 87–112.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research*, 29(9), p.e45.

Pinto, P. a B. et al., 2011. RNA polymerase II kinetics in polo polyadenylation signal selection. *The EMBO journal*, 30(12), pp.2431–44.

Poort, S.R. et al., 1996. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood*, 88(10), pp.3698–703.

Proudfoot, N., 2004. New perspectives on connecting messenger RNA 3' end formation to transcription. *Current opinion in cell biology*, 16(3), pp.272–8.

Proudfoot, N.J. & Brownlee, G.G., 1976. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature*, 263(5574), pp.211–4.

Rehfeld, A. et al., 2013. Alterations in polyadenylation and its implications for endocrine disease. *Frontiers in endocrinology*, 4(May), p.53.

Rehwinkel, J. et al., 2005. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA*, 11(10), pp.1530–44.

Resch, A. et al., 2004. Evidence for a subpopulation of conserved alternative splicing events under selection pressure for protein reading frame preservation. *Nucleic acids research*, 32(4), pp.1261–9.

Rio, D.C. et al., 2010. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harbor protocols*, 2010(6), p.pdb.prot5439.

Roberts, G.C. et al., 1998. Co-transcriptional commitment to alternative splice site selection. *Nucleic acids research*, 26(24), pp.5568–72.

Rocha, J.J.E. et al., 2011. A phagocytic route for uptake of double-stranded RNA in RNAi. *PloS one*, 6(4), p.e19087.

Sanchez, S.E. et al., 2011. Alternative splicing at the right time. *RNA biology*, 8(6), pp.954–9.

Sandberg, R. et al., 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science (New York, N.Y.)*, 320(5883), pp.1643–7.

Saunders, A. et al., 2003. Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. *Science*, 301(5636), pp.1094–6.

Saunders, A., Core, L.J. & Lis, J.T., 2006. Breaking barriers to transcription elongation. *Nature reviews. Molecular cell biology*, 7(8), pp.557–67.

Schneider, I., 1972. Cell lines derived from late embryonic stages of Drosophila melanogaster. *Journal of embryology and experimental morphology*, 27(2), pp.353–65.

Shandilya, J. & Roberts, S.G.E., 2012. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochimica et biophysica acta*, 1819(5), pp.391–400.

Sheets, M.D., Ogg, S.C. & Wickens, M.P., 1990. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic acids research*, 18(19), pp.5799–805.

Shepard, P.J. et al., 2011. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA (New York, N.Y.)*, 17(4), pp.761–72.

Shi, Y., 2012. Alternative polyadenylation: new insights from global analyses. *RNA (New York, N.Y.)*, 18(12), pp.2105–17.

Shilatifard, A. et al., 1996. An RNA polymerase II elongation factor encoded by the human ELL gene. *Science*, 271(5257), pp.1873–6.

Shopland, L.S. et al., 1995. HSF access to heat shock elements in vivo depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. *Genes & Development*, 9(22), pp.2756–2769.

Smibert, P. et al., 2012. Global patterns of tissue-specific alternative polyadenylation in Drosophila. *Cell reports*, 1(3), pp.277–89.

Smith, E., Lin, C., Garrett, A.S., et al., 2011. The little elongation complex regulates small nuclear RNA transcription. *Molecular cell*, 44(6), pp.954–65.

Smith, E., Lin, C. & Shilatifard, A., 2011. The super elongation complex (SEC) and MLL in development and disease. *Genes & development*, 25(7), pp.661–72.

Smith, E. & Shilatifard, A., 2013. Transcriptional elongation checkpoint control in development and disease. , pp.1079–1088.

Sunkel, C.E. & Glover, D.M., 1988. polo, a mitotic mutant of Drosophila displaying abnormal spindle poles. *Journal of cell science*, 89 (Pt 1), pp.25–38.

Svejstrup, J.Q., 2013. RNA polymerase II transcript elongation. *Biochimica et biophysica acta*, 1829(1), p.1.

Svetlov, V. & Nudler, E., 2013. Basic mechanism of transcription by RNA polymerase II. *Biochimica et biophysica acta*, 1829(1), pp.20–8.

Takagaki, Y. et al., 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell*, 87(5), pp.941–52.

Takahashi, H. et al., 2011. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell*, 146(1), pp.92–104.

Tan, S., Conoway, R.C. & Conoway J.W., 1995. Dissection transcription required elongation. *Proceedings of the National Academy of Sciences*, 92(June), pp.6042–6046.

Thirman, M.J. et al., 1994. Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 91(25), pp.12110–4.

Tian, B. et al., 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic acids research*, 33(1), pp.201–12.

Tian, B. & Manley, J.L., 2013. Alternative cleavage and polyadenylation: the long and short of it. *Trends in biochemical sciences*, 38(6), pp.312–20.

Wahle, E., 1991. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*, 66(4), pp.759–68.

Wahle, E. & Keller, W., 1996. The biochemistry of polyadenylation. *Trends in biochemical sciences*, 21(7), pp.247–50.

Weaver, R.F., 2012. *Molecular biology.*, New York: McGraw-Hill.

Wilusz, J., Pettine, S.M. & Shenk, T., 1989. Functional analysis of point mutations in the AAUAAA motif of the SV40 late polyadenylation signal. *Nucleic acids research*, 17(10), pp.3899–908.

Winnebeck, E.C., Millar, C.D. & Warman, G.R., 2010. Why does insect RNA look degraded? *Journal of insect science*, 10, p.159.

Wood, A.J. et al., 2008. Regulation of alternative polyadenylation by genomic imprinting. *Genes & Development*, 22, pp.1141–1146.

Yamada, T. et al., 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Molecular cell*, 21(2), pp.227–37.

Yamaguchi, Y. et al., 1999. NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. *Cell*, 97(1), pp.41–51.

Yang, Q. & Doublié, S., 2011. Structural biology of poly(A) site definition. *Wiley interdisciplinary reviews. RNA*, 2(5), pp.732–47.

Yoh, S.M. et al., 2007. The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes & development*, 21(2), pp.160–74.

Zanoni, I., Ostuni, R. & Granucci, F., 2009. Drosophila Schneider 2 (S2) cell culture. *Protocol Exchange*.

Zhao, J. & Hyman, L., 1999. Formation of mRNA 3 J Ends in Eukaryotes : Mechanism , Regulation , and Interrelationships with Other Steps in mRNA Synthesis. *Microbiology and Molecular Biology Reviews*, 63(2), pp.405–445.

Zhou, J., Schweikhard, V. & Block, S.M., 2013. Single-molecule studies of RNAPII elongation. *Biochimica et biophysica acta*, 1829(1), pp.29–38.

Appendix

Table A1- Oligonucleotides used during the experimental work for PCR and qPCR. All primers were obtained from Sigma- Aldrich.

		Sequence (5'- 3')
polo total	*Forward*	CCGTACAACATGTGCCGTAG
	Reverse	CTTTAGACACGCCGTTCTCC
polo pA2	*Forward*	ACGTGTTTCGAAATGCCTAT
	Reverse	ACACTTAAACACTTTGCAGCAG
CG6024 total	*Forward*	CACCGCACTCCACACAACTA
	Reverse	ATTGGGATGTCCGGTTCC
CG024 pA2	*Forward*	CAGTAACGGAAGACCCGAAA
	Reverse	GGTCCAAAGGAGGGTGAAAT
7SL	*Forward*	TTGGCTAAGGAGGGATGAAC
	Reverse	CTACTGCCTACCACGGGAAC
dELL	*Forward*	TGTGGAACGATGTGGACGAG
	Reverse	GCGGTGTCAGATTTTGAGGC
dSpt6	*Forward*	GGCCGTCTCCGATAGTAGC
	Reverse	TCGATCAGATCTTTGAGCTCTTC

Table A2- Oligonucleotides used to obtain the DNA templates for the production of dsRNA

		Sequence (5'- 3')
T7- dELL	*Forward*	TAATACGACTCACTATAGGGAGAGCGTCAAAGAAGTGCCAGTG
	Reverse	TAATACGACTCACTATAGGGAGACGTAGTCCCCGTATCCGTTG
T7- dSpt6	*Forward*	TAATACGACTCACTATAGGGAGACCGTAACCCCGGTGCCCGAGG
	Reverse	TAATACGACTCACTATAGGGAGAGGCTCTTGTGCCAGCTGTCGG
T7- DsRED	*Forward*	GAATTAATACGACTCACTATAGGGAGACTTCAAGGTGCGCATGGAG
	Reverse	AATTAATACGACTCACTATAGGGAGAGGACTTGAACTCCACCAGGTAGTG

Table A3- PCR programs used to obtain the DNA templates for the production of dsRNA

PCR product	PCR reaction				
	Initial denaturation step (ºC)	Denaturation temp (ºC) (step 2)	Annealing temp (ºC) (step 3)	Extension temp (ºC) (step 4)	Nº of cycles (step 2-4)
dELL	95(5min)	95 (30 sec)	53 (30 sec)	72 (1 min)	3
		95 (30 sec)	55 (30 sec)	72 (1 min)	5
		95 (30 sec)	57 (30 sec)	72 (1 min)	27
dSpt6	95(5min)	95 (30 sec)	55 (30 sec)	72 (1 min)	3
	95(5min)	95 (30 sec)	57 (30 sec)	72 (1 min)	5
	95(5min)	95 (30 sec)	60 (30 sec)	72 (1 min)	27
DsRED	95(5min)	95(1min)	55 (30 sec)	72 (1 min)	**35**